

SCOMM

137:26



HOUSE SPECIAL COMMITTEE ON EDUCATION

STATE CAPITOL, ROOM 106

COMMITTEE MEMBERS

Rep. Mark Neuman
Chairman
Room 432
465-3820

Rep. Carl Gatto
Vice-Chair
Room 411
465-3743

Rep. Bob Lynn
Room 415
465-4931

Rep. Bill Thomas
Room 428
465-3732

Rep. Peggy Wilson
Room 108
465-3824

Rep. Les Gara
Room 418
465-2647

Rep. Woodie Salmon
Room 114
465-4527

MEMORANDUM

Date: March 10, 2006

To: Members of House Special Committee on Education

From: Representative Mark Neuman, Chairman
House Special Committee on Education

Re: Education Committee Schedule

Tuesday, March 14, 2006, at 11:00 a.m. Room 106

HB 247 Class Size Reduction Grants
Representative Croft



REPRESENTATIVE ERIC CROFT

Memorandum

TO: Representative Mark Neuman, Chairman Special Committee on Education

FROM: Representative Eric Croft

DATE: April 6, 2005

RE: Hearing Request, HB 247

Please schedule HB 247 for a hearing in the Special Committee on Education at your earliest convenience. I have attached a copy of the bill, a sponsor statement and some additional background information. We have been assured of support from a variety of groups statewide, including teacher, parent, and social welfare organizations, and I will submit their written statements to you as my office receives them. I appreciate your help in moving this important legislation through the committee process.

Thank you.





REPRESENTATIVE ERIC CROFT

Sponsor Statement HB 247 The Time to Teach Act of 2005

Reduced class sizes are a priority for parents and leaders in this state. The Time to Teach Act proposes an evidence-based program that would dramatically increase the academic performance of Alaskan students by reducing class sizes in K-3 to 15 or less students.

Research shows that outcomes for students do not significantly improve until class sizes are dramatically reduced, especially in the early years. Specifically, Project STAR, a landmark study conducted in Tennessee, suggests that students in K-3 who were taught in classrooms of 15 or less students not only perform significantly better on average than cohorts in larger classes, but also continue to perform better as they move into more advanced grades and larger classrooms. Also of note are findings that suggest that minority students especially benefit from reduced class sizes.

Because reducing class size is so important in the primary grades, school districts would benefit more from funding dramatic decreases in kindergarten through 3rd grades than from reducing class size by a few students across the board. Under the Time to Teach Act, any school that commits to reducing class sizes to 15 or fewer students for grades K-3 would be eligible for a grant from the state. Grants would be administered school by school with each school committing to studying the academic progress of the students from smaller classes.

Allowing districts to decide when and how to reduce class sizes prevents problems that could be caused by a law that mandates that all schools make reductions immediately. Because of the demands of hiring new teachers and creating space for new classrooms, other states have created problems for school districts by suddenly mandating reduced class sizes. Schools should be able to work with parents and their communities to decide how and when they can take this step. The Time to Teach Act allows for this local control and flexibility.

Finally, the Time to Teach requires schools to study the academic progress of students in reduced class sizes for at least five years. This study will not only allow the state to evaluate the effectiveness of the program, it will also provide information to stakeholders interested in improving or expanding the program over time.



The Time to Teach Act - HB 247
Subject Sectional

Section 1. Short Title

Section 2.

Sec. 14.17.485(a). Establishes a grant program within the base student allocation (BSA) foundation formula to reduce class sizes. The program would allow any school that voluntarily elects to set K-3 class size at no more than 15 to receive an increased base allocation of \$8000 for each student in K-3 classes. The proposed BSA of \$8000 would be adjusted annually for inflation.

Sec. 14.17.485(b). The grant program would require each eligible school to enter into an agreement with their respective districts. The agreement would require the school to commit to the following:

- (1) The school will maintain reduced class sizes for K-3 for at least 5 years.
- (2) The school will complete a study that compares the academic progress of students benefiting from reduced class sizes to students without the benefit of reduced class size.
- (3) In the reduced size classrooms, the school will agree to place only teachers with valid certificates, qualified to teach in an area of their expertise. If children with identified disabilities are placed in reduced size classroom, then at least one special education aide will be assigned to that classroom.

Sec. 14.17.485(c). Because student with disabilities can take significantly more time to teach, for the purposes of calculating class size, a student with an identified disability will count as two students.

Sec. 14.17.485(d). If a school enters into the agreement with the district to reduce class sizes and fails to comply with or enforce the requirements of this section, then the department or a parent of a child enrolled in the school can file a complaint in a court of relevant jurisdiction.

Sec. 14.17.485(e). If there is not enough money to meet the amounts authorized under bill then the department will reduce grant funding pro rata.

Sec. 14.17.485(f) "Class size" means a group of students taught in a regular by *one teacher*. A classroom of over 15 students taught by *two teachers* would not count a reduced sized classroom.

Small Classes in American Schools: Research, Practice, And Politics

Mr. Finn provides a brief overview of the current status of class-size reduction programs in the U.S., summarizes the research base that has moved districts and states to seek class-size reduction, and calls attention to the misapplication of the research in some contexts. He also discusses questions about smaller classes that remain unanswered and describes current research into the long-term consequences of small classes and efforts to explain why they are effective.

BY JEREMY D. FINN

IN RECENT YEARS, more than half of the states, countless districts, and the federal government have sponsored class-size reduction (CSR) programs. In California alone 28,000 new teachers were hired in the first three years of the statewide CSR initiative. Meanwhile, as part of its effort to reduce class sizes, the U.S. Department of Education has paid the salaries of 29,000 new teachers, mainly in poor urban school districts. Nationwide, it is impossible to count the number of new teachers hired and the number of classes that have been reduced in size. But these numbers are certainly large.

JEREMY D. FINN is a professor in the Graduate School of Education, State University of New York, Buffalo. He wishes to acknowledge that the ideas expressed here have been influenced heavily by his collaborations with Charles Achilles and Gina Panno. Work on this article was supported in part by grants from the Spencer Foundation and the William T. Grant Foundation, though the opinions expressed are the author's own. An earlier version of this article was presented to the Mid-Western Educational Research Association in Chicago, October 2001.



Why is it that, after years of research and debate but little action, class sizes are finally being reduced in the elementary grades across the U.S.? There are a number of reasons, including the following:

- Everybody likes the idea of small classes — teachers, parents, policy makers, legislators, and even the courts¹ understand the importance of small classes for teaching and learning;
- High-quality research has demonstrated the benefits of small classes in the early grades — especially for students at risk;

- Until very recently education had risen to the top of state and national agendas; and

- The economy was healthy, so we had ample resources to direct toward school improvement.

But much has changed in recent months, making the future of smaller classes in the elementary grades less clear. The White House education plan, "No Child Left Behind," earmarked the federal reduced-class-size initiative as one of two programs to be eliminated. The recent instability in the economy may have left states and districts less able to hire additional teachers. And the events of September 11 refocused our national attention in a way that may well give lower priority to education issues. It remains to be seen if small class sizes have become sufficiently important and sufficiently institutionalized that they will continue to be part of our basic educational plans.

In this article I provide a brief overview of the current status of class-size reduction programs in the U.S.; the overview is brief because the situation is changing even as I write. I also summarize the research base that has moved districts and states to seek class-size reduction and call attention to the misapplication of the research in some contexts. Finally, I discuss questions about smaller classes that remain unanswered and describe current research into the long-term consequences of small classes and efforts to explain why small classes are effective.

My basic premise is that small classes in the elementary grades are *not* the solution to the problems of American schools; they are not the "silver bullet." Instead, small classes provide an *essential opportunity* for instruction to be more effective and for students to become optimally involved in the learning process. The most important question yet to be answered is "How can educators take best advantage of small class sizes to maximize learning outcomes?"

THE RESEARCH BASE

Prior to the 1980s, dozens of studies were conducted on the relationship between class size and pupil performance. Many of these suffered from methodological flaws, including small samples, poor research designs, and inadequate treatment of the data. Not one was a large-scale randomized experiment. However, reviews of this research found consistencies and supported some tentative conclusions:

- Reduced class size (below 20 pupils) can be expected to produce a modest increase in academic achieve-

ment.²

- Small classes are most beneficial in the early primary grades.³

- Students who are economically disadvantaged are most likely to benefit from small classes.⁴

PROJECT STAR

Beginning in 1985, the state of Tennessee, under Gov. Lamar Alexander, undertook a remarkable experiment to test these propositions. Project STAR (Student/Teacher Achievement Ratio) was a large-scale randomized within-school experiment, most unusual in education. Pupils entering kindergarten in each participating school were assigned at random to a small class (13-17 students), a full-size class (22-26 students), or a full-size class with a full-time teacher aide. Teachers were also assigned at random to the classrooms. Pupils were kept in the same condition — small, regular, or teacher aide — for up to four years (grade 3), with a new teacher assigned at random each year.

In addition to being a scientifically controlled experiment, STAR had other special features:

- The study was extensive. More than 6,000 students in 329 classrooms in 79 schools participated in the first year, and almost 12,000 students participated at some point in the four-year study.

- The class arrangement was maintained all day, all year long, and there was no other intervention. For example, there was no special training for teachers and no special curricula.

- Both norm-referenced and criterion-referenced achievement tests were administered to each pupil in the spring of each year. Other data on the students (e.g., classroom behavior, retentions in grade) and on their teachers were collected regularly.

- STAR pupils were followed after they returned to full-size classes in grade 4. Measures of academic performance and other outcome data were collected through high school, and information on postsecondary schooling and employment is still being collected.

The findings of Project STAR are important building blocks in today's knowledge base about small classes. From a scientific perspective, it is also important that the STAR data have been reanalyzed by researchers not associated with the project who have used a variety of statistical procedures.⁵ With only minor exceptions, these analyses have agreed with the original findings. Because STAR is one of the few large-scale randomized experiments in the field of education, it

has also been used by other agencies as a model in the push for more experiments to support education policies.

The findings of STAR have been summarized in a number of publications.⁶ The most intensive look at short- and long-term academic outcomes to date appeared in *Teachers College Record* in 2001.⁷ In that article, effect sizes are presented in two metrics, the usual "standard deviation" scale and in terms of "months of schooling." In the sections that follow, I summarize the highlights of the findings of Project STAR, drawn from those reports.

TABLE 1.
Small-Class Advantage in Months of Schooling, K-3*

Grade	Stanford Achievement Test Scale		
	Total Reading	Word Study Skills	Total Mathematics
Kindergarten	0.5	0.5	1.6
Grade 1	1.3	0.9	2.8
Grade 2	3.9	4.8	3.5
Grade 3	4.5	5.5	2.6

*Difference between mean of small classes and mean of full-size classes on grade-equivalent scale. Adapted from Table 2 of Jeremy D. Finn et al., "The Enduring Effects of Small Classes," *Teachers College Record*, vol. 103, 2001, p. 154.

Experimental years (K-3). Small classes had statistically significant academic benefits in every grade, in *all academic subjects*. The differences between small classes and full-size classes were in the range of 0.2 to 0.3 standard deviations in each school subject. Table 1 shows the advantage of attending small classes in terms of "months of schooling." For example, kindergarten students who attended small classes were about one-half month ahead of students who attended larger classes in reading and about 1.6 months ahead in mathematics. In grade 1, the small-class advantage was 1.3 months in reading and 2.8 months in mathematics. By spring of grade 3, the advantage in reading was 4.5 months (or about half a school year).

The magnitude of effects was greater for students who *started early* and *spent more years* in small classes. For example, first-graders who were in small classes for the first time were about 0.6 of a month ahead of their schoolmates in reading and about two months ahead in mathematics. First-graders who were in small classes for the second year (since kindergarten) were about two months ahead

of their schoolmates in reading and about 3.5 months ahead in mathematics. A similar pattern was found for students in grades 2 and 3 on every test.

In every grade, the benefits of small classes were *greater for minority students or students attending inner-city schools* than for white students in nonurban schools. Statistically, the advantages were often as much as two to three times as great, thus reducing the white/minority achievement gap. For example, the racial gap in pass rates on the grade 1 reading mastery test was 14.3% in full-size classes—that is, 14.3% more white students than black students mastered the reading objectives.

In small classes, the gap was reduced to 4.1%. Both black and white students gained significantly by being in small classes, but black students gained more. Economist Alan Krueger reanalyzed the STAR data and described the reduction in the achievement gap in other terms. He concluded that, by third grade, the black/white gap in school performance would be reduced by 38% if all students attended small classes.⁸

Post-experimental years (4-8). After all students returned to full-size classes, the benefits of small classes continued to be statistically significant through all subsequent grades in all subject areas.⁹ Table 2 shows the advantage in months of schooling in grades 4-8 of having attended small classes in grades K-3. The benefits in grade 4 ranged from 2.4 months to 4.6 months; in grade 6, from 3.0 to 5.1 months (apart from science); and in grade 8, from 3.4 months to 4.8 months.¹⁰ (The values in Table 2 are conservative because they include students who attended small classes for one, two, three, and four years.)

Starting early and continuing in small classes for multiple years are both related to the strength and duration of carryover effects. In any given grade (4-8), the

TABLE 2.
Overall Small-Class Advantages for Grades 4, 6, and 8*

Grade	CTBS Scale			
	Total Reading	Total Mathematics	Science	Social Science
Grade 4	3.9	2.4	3.1	4.6
Grade 6	3.0	3.1	1.0	5.1
Grade 8	4.1	3.4	4.3	4.8

*Difference between mean of small classes and mean of full-size classes on grade-equivalent scale. Adapted from Table 4 of Finn et al., p. 167.

effects were larger for students who had spent more years in small classes in grades K-3. For example, at the end of grade 6, students who had attended small classes for one year had a 1.2-month advantage in reading over students who had attended full-size classes. Students who had attended small classes for two years had a 2.8-month advantage; three years in a small class produced a 4.4-month advantage. And so on, in each school subject at every grade level.

OTHER CSR INITIATIVES AND THEIR FINDINGS

A large number of (nonexperimental) CSR initiatives have been undertaken since Project STAR, but many have not been systematically evaluated. Those that have been evaluated include Wisconsin's Project SAGE; the well-researched effort in Burke County, North Carolina; and California's statewide CSR program. The Wisconsin and North Carolina projects targeted schools serving low-income students. While all these programs have unique features, all have yielded findings consistent with STAR.

Project SAGE. Wisconsin's Student Achievement Guarantee in Education (SAGE) program, begun in 1996, was targeted to schools in which 50% or more of the students were eligible for Title I support. Participating schools were required to "systematically reduce class size to 15 in grades K-3; stay open from early in the morning to late in the day, . . . provide a rigorous academic curriculum to improve academic achievement, and establish staff development and accountability mechanisms."¹¹ In its first year of operation, SAGE was implemented in 30 schools in 21 districts across Wisconsin; it has expanded considerably since then. The majority of SAGE classrooms were self-contained small classes. Schools with insufficient space implemented alternative classroom models, including team-taught classrooms with two teachers and approximately 30 students and "floating teacher" classrooms, in which an additional teacher was brought in to share instructional responsibilities during reading and mathematics lessons.

An evaluation was conducted in each of four years, comparing SAGE classes to a matched comparison group of non-SAGE classes in the same districts. The results show consistent benefits of small classes with effect sizes similar to those of STAR.¹² The small-class advantage was also found to be significantly greater for minority students than for white students.

Burke County, North Carolina. Burke County is a

relatively poor rural community with an increasingly diverse population. Thirty-four percent of students in the Burke County schools are eligible for free or reduced-price lunch; 11% of the students have limited proficiency in English. In 1991, four Burke County elementary schools were selected to implement small classes (no more than 15 students) in grade 1. The following year the initiative was expanded to include grade 1 in 14 elementary schools and grade 2 in the original four schools. By 1995-96, small classes had been implemented in grades 1 and 2 in all schools and in grade 3 in six schools. Intensive staff development programs accompanied reduction in class size. By reallocating existing staff members, the schools in Burke County reduced class sizes with no increase in per-pupil expenditures.

The evaluation of the program included extensive classroom observations as well as statistical comparisons of small classes with matched full-size classes on reading and mathematics tests. Students in smaller classes had higher achievement scores in reading and mathematics in grades 1 through 3, and they continued to have higher reading scores in grades 4 through 7, after returning to full-size classes.¹³

California's statewide program. In the spring of 1996, the California legislature passed a reform initiative to limit class sizes statewide in grades K-3 to a maximum of 20 students. By the time school began in the fall of 1996 — just a few months later — 12,000 new teachers had been hired, and class sizes were reduced in 14% of the state's kindergarten classrooms, in 88% of its grade-1 classrooms, in 57% of its grade-2 classrooms, and in 18% of its grade-3 classrooms. By 1998-99, class sizes were reduced in virtually all grade-1 and grade-2 classrooms across the state and in 84% to 86% of kindergarten and grade-3 classrooms.

Initially, the state did not plan a major evaluation of its CSR initiative, despite the fact that it cost over \$1 billion per year. A broad-based evaluation was organized by a consortium of organizations, with funding from private and public sources.¹⁴ The evaluation of academic achievement was limited largely to grade 3; since most classes in grades 1 and 2 were reduced simultaneously, no comparison groups were available. In grade 3, students in small classes outperformed their peers who attended large classes to a small, but statistically significant, degree.¹⁵ Effect sizes in reading and mathematics ranged from 0.05 to 0.10 standard deviations. These effect sizes are very close to the effects found by Project STAR for grade-3 students with one

or two years of small-class participation. Larger effects were found in STAR for students who began in kindergarten or grade 1.

TEACHER AIDES

Project STAR's results for teacher aides have largely been overlooked because of the excitement about the findings for small classes, but the policy implications are considerable. When STAR was designed, Tennessee policy makers wondered if teacher aides could provide the same benefits as small classes but at a substantially lower cost. But the STAR analyses continually reported "no significant difference" between classes with teacher aides and full-size classes without aides. Those results were assembled and extended in several recent reports.¹⁶

Susan Gerber and her colleagues have estimated that there were more than 600,000 teacher aides in American classrooms in 1998, at an annual cost of about \$9 billion.¹⁷ We would certainly hope for demonstrable benefits from such an expenditure. Unfortunately, virtually all the research on the topic, including Project STAR, has found that, in general, teacher aides benefit neither teachers nor students.

To illustrate, consider some findings from a study reported in 2000 by my colleagues and me.¹⁸ In that report, the academic performance of students in classes with teacher aides was compared to both other class types (small and full-size classes without aides), ratings of student behavior were compared across the class types, and teachers in the three class types reported the severity of problems they encountered in the classroom in terms of time management, managing and controlling the class, and engaging students in productive learning activities.

The study posed two questions. The first was "Do students in teacher aide classes perform as well or better as well as do students in small classes?" To quote from the report:

The answer is unequivocally "no." In terms of academic achievement, students in small classes performed significantly better on every test administered in every grade. There were no exceptions. . . . In terms of behavior . . . students who had attended small classes exhibited superior learning behaviors on two of three dimensions and on total engagement (in learning). . . . These attributes — academic achievement and engagement behaviors — are so

fundamental that on this basis alone, it is clear that teaching assistants are not a substitute for the benefits provided by reduced class size. When teachers were interviewed about their preference, 71% said they would prefer teaching a small class to teaching a regular class with a full-time assistant.¹⁹

The second question was "Do classes with teaching assistants have advantages over full-size classes without assistants?" The results led to these conclusions:

Here, too, the answer is "no." No overall differences in academic achievement were found between the performance of students in teacher aide classes and students in regular classes on any test in any grade. . . . In several instances, students in aide classes performed more poorly than did students in non-aide classes. . . . In terms of learning behavior, again no significant differences were found . . . in Grade 4 or Grade 8. In several instances, behavior was marginally poorer among students in classes with aides.²⁰

Finally, teachers with aides reported little or no relief from the responsibilities of teaching, even when teaching assistants were classified according to the types of duties they performed: administrative, noninstructional interactions with students, or instruction.

Some districts (e.g., Burke County and San Diego) have used teacher aide funds to hire additional teachers. Given the absence of positive impacts and even the possibility of negative effects of aides, this seems to be a sensible policy. The other option discussed in the reports — to "remedy the deficient preparation of paraprofessionals for the tasks they perform, the lack of clearly defined roles for aides in the classroom, and the absence of training for teachers in utilizing their assistants"²¹ — also deserves serious consideration.

OTHER FINDINGS ABOUT SMALL CLASSES

Project STAR did not undertake sufficient studies of classroom processes. However, between the limited process research undertaken in STAR and research conducted in conjunction with other CSR initiatives, an array of additional findings has emerged, among them:

- teacher morale is improved in small classes;²²
- teachers spend more time on direct instruction and less on classroom management when classes are smaller;²³
- there are fewer disruptions in small classes and fewer

discipline problems;²⁴

- students' engagement in learning is increased;²⁵
- in-grade retentions are reduced;²⁶
- dropout rates may be reduced;²⁷ and
- greater numbers of students who attend small classes in the early grades elect to take SAT or ACT tests in high school — that is, aspirations to attend college are increased, especially among African American students.²⁸

THE COSTS OF REDUCING CLASS SIZES

Policy makers often want to know if the benefits produced by small classes are worth the costs, because small classes have been described as an expensive intervention. Unfortunately, only limited efforts have been made to answer this question. The most complete analysis of costs was performed by Dominic Brewer and his colleagues, who estimated the nationwide costs of CSR under such policy alternatives as small classes of different sizes, different amounts of spending flexibility, and targeting small classes to students at risk.²⁹ Indeed, this work produces some very large cost figures — \$2 billion and up.

However, none of these analyses have looked at possible resource tradeoffs, nor have they examined any factors on the benefit side of the equation. The issue of tradeoffs is quite complex, so it is not surprising that it hasn't been studied. Several districts used funds allocated for hiring teacher aides to hire additional teachers and so reduce class sizes. In financially poor Burke County, small classes were implemented in grades 1 through 3 by reassigning existing staff members with no increase in per-pupil expenditures. Researchers should give more attention to the approaches taken in these districts, especially if resources become less plentiful.

Nevertheless, the omission of benefits and the focus on costs alone in these analyses is surprising. The only analysis of benefits in recent years may be that produced by economist Alan Krueger. Using data from STAR and previous studies of school performance and personal income, Krueger concluded that the benefits — in terms of students' future earnings — of reducing class sizes are very close to the per-pupil cost of smaller classes. In other words, the costs are recovered in the form of personal income to the students.³⁰

However, the total benefits are greater still. If small classes are an incentive for teachers to remain in urban settings, if students are more likely to attend college, and if grade retentions and dropout rates are decreased, then these outcomes represent benefits as well. Such

data need to be included in any complete analysis of the costs and benefits of small classes.

HOW TO DO SMALL CLASSES THE WRONG WAY

The widespread implementation of reduced class sizes has produced some "tried and true" ways to negate the benefits. Policy makers and administrators should be aware of two pitfalls in particular.

OVERLOOKING THE NEED FOR PROFESSIONAL SUPPORT AND DEVELOPMENT

In the rush to hire and place new teachers in classrooms, it is tempting for administrators to overlook the need for professional development and support. The California CSR initiative demonstrated the serious side effects of doing things too quickly. In its haste to reduce K-3 class sizes in a matter of a few months, the state placed many individuals in classrooms who had less than full teaching credentials and lacked adequate experience managing students. The effect was so large that the preparation level of the entire state's teaching force declined.³¹ Since districts serving low-income students, minority students, or students learning English had more difficulty attracting qualified teachers, the decline in teacher qualifications was greatest in those districts. For example, schools serving 30% or more low-income students saw a 16% drop in fully credentialed teachers between 1995-96 and 1998-99. Similar problems have occurred in many other states and districts.

When inexperienced teachers are placed in classrooms without adequate preparation, the result can be a level of disorganization that requires weeks or even months to remedy. These dynamics can easily offset the academic benefits of small classes. Brian Stecher and his colleagues recommend that CSR initiatives be undertaken slowly and with careful planning.³² Well-designed programs of professional support and development can also help.

Project STAR demonstrated that the benefits of small classes are obtained *without* any special teacher preparation. However, STAR had an abundant supply of credentialed teachers — a resource not always available today. Several major CSR programs — for example, in Burke County, North Carolina; in Fairfax County, Virginia; and Wisconsin's Project SAGE — have made effective use of professional development in con-

junction with reduced classes. Experience shows how professional support and development activities can be put to best use.

- Many teachers placed in elementary classrooms are new to teaching, new to the classroom, and new to their school setting. They need help "getting started," and they need mentoring and other forms of support throughout their first year of teaching.

- Many veteran teachers are transferring from other settings to small classes. The instructional practices they have adopted over the years are not always "current best practice." They can benefit from updating and relearning those skills that work well with small classes but were too arduous to practice in large classes.

- It may be possible to *enhance* the benefits of small classes by taking advantage of the opportunities they provide. For example, professional development can show teachers how to cover content in greater depth and how best to take advantage of the increased sense of community that is typical of small classes.³¹

Not all forms of professional development can accomplish these purposes effectively. Reduced class size creates its own particular needs. Support and development activities must focus on the specific needs of small-class teachers and follow the principles of effective programming.³²

CONFUSING PUPIL/TEACHER RATIOS WITH CLASS SIZE

The confusion of these two terms is common and constitutes a second major pitfall to be avoided in implementing CSR initiatives. "Class size" is the number of students regularly in a teacher's classroom for whom that teacher is responsible. The idea of class size is especially important to *teachers* because it constrains all their interactions with pupils, encourages or discourages learning behavior and pro- or antisocial behavior, and is clearly related to the amount of material students learn. As Charles Achilles puts it, "A class with 15 students and one teacher has a class size of 15. A class with 28 pupils and one teacher has a class size of 28. A class with 28 pupils and two teachers and a full-time teacher aide still has a class size of 28."³³

The "pupil/teacher ratio" is the ratio of the number of students in an educational unit to the number of full-time-equivalent professionals assigned to that unit. Pupil/teacher ratios have been used by *economists* for many years to develop funding formulas for dis-

tricts and states. However, the pupil/teacher ratio for a school, district, state, or nation *does not describe the proximal setting in which pupils are learning*. In the U.S., many large urban districts have small pupil/teacher ratios (e.g., 13 to 1) because of the large number of ancillary staff members, even though most students spend the entire school day in overcrowded classrooms.³⁶

This distinction between class size and pupil/teacher ratio is important for two reasons. First, the strong research base on small classes does not apply to large classes — no matter how many teachers are present. Some schools, facing a shortage of classroom space, have created large classes with several teachers or with teachers and aides instead of small classes. Indiana's Prime Time, which purported to be a CSR program, had class sizes ranging from 12 to 31; many "small" classes were actually large classes with a teacher and an aide. Not surprisingly, the achievement results of Prime Time have been quite mixed.³⁷

The possibility of having two full-time teachers "team-teach" a large class is permitted under the regulations guiding the U.S. Department of Education's CSR initiative, and this model has been used in some Project SAGE classrooms. Although the SAGE evaluation suggests that team-taught classes are sometimes as effective as small classes,³⁸ it has not been confirmed with large samples or through a controlled experiment. Simply put, we don't have the same level of scientific information about how these other classroom arrangements work.

The second reason that maintaining the distinction between small classes and low pupil/teacher ratios is important is that critics have used data on pupil/teacher ratios to attempt to show that small classes are not beneficial.³⁹ Because pupil/teacher ratios are usually computed for large, heterogeneous units (i.e., school districts, states, or countries), it is little surprise that they have a weak relationship with academic achievement. These levels of analysis may be appropriate for an economist's work, but they are not useful for educators concerned with teaching and learning in individual classrooms.⁴⁰

UNANSWERED QUESTIONS AND ONGOING RESEARCH

Many questions remain to be answered with regard to small classes. Among those that are asked repeatedly, we find:

- How small is "small"? Is a class of 20 students like-

ly to be as effective as, say, a class of 17?

- How effective are small classes in the middle grades? In high school?

- Can the effects of small classes be enhanced through particular instructional strategies? By combining CSR with other interventions, such as preschool programs or remedial programs? By taking advantage of the improved sense of community that arises in small classes?

Two broad questions are being addressed currently. First, what are the long-term impacts of small classes in the early grades? Princeton University's Alan Krueger and his colleagues are augmenting the STAR data with information about students who take college admissions tests, information about child bearing, and information about delinquent or criminal behavior. They will also collect information about unemployment rates. Preliminary reports have already shown that class size is associated with some of these outcomes."

Also, together with Jayne Boyd-Zaharias of HEROS, Inc., the primary organization for Tennessee's STAR research, I am performing additional follow-up analyses. Among them will be an examination of the high school courses that STAR participants have taken, their high school grades, and their graduation rates. Approximately 500 telephone interviews are also being conducted with former STAR participants to document postsecondary schooling and employment. In all, this will create a formidable 17-year database for continuing research.

The second question I am pursuing, along with

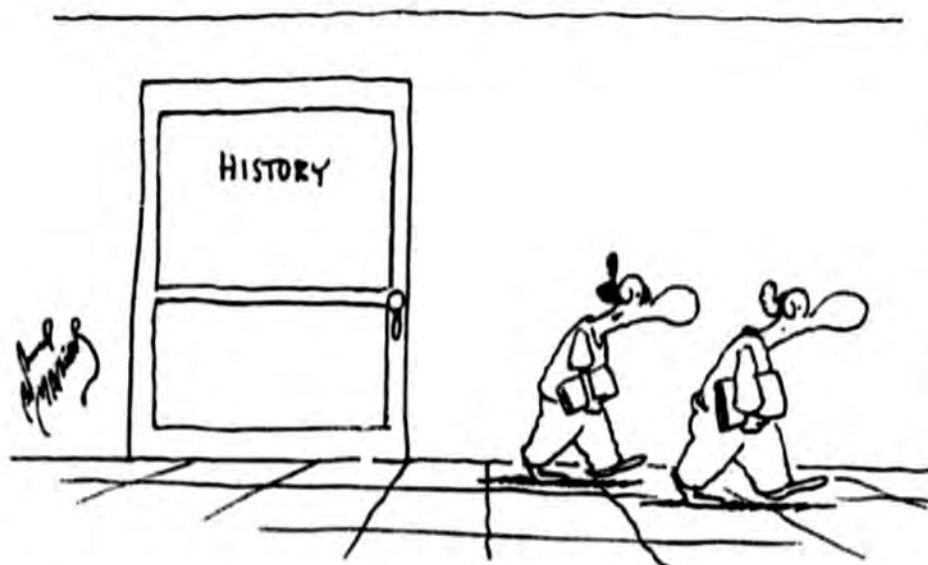
Gina Pannozzo and Charles Achilles, is the "black box" question: Why do small classes work as well as they do? People often speculate that teachers change their instructional styles in small classes, providing more one-on-one teaching and higher-quality instruction. Interestingly, neither STAR nor other process studies support this hypothesis. It is fairly clear that teachers of small classes spend somewhat more time on direct instruction and somewhat less time on classroom management and discipline. Beyond that, few if any qualitative differences occur when class sizes are reduced. Changes in instruction are generally small and do not explain the consistent academic benefits that are found.

This new line of research focuses on a second hypothesis: that students become better students in small classes. That is, students become more engaged in learning, and they display more prosocial behavior and less antisocial behavior. Past research on the topic consists of 15 published studies of students' learning and social behavior. These studies vary considerably in quality. Nevertheless, of 46 measures of students' engagement in learning, 30 are consistent with the hypothesis, and not one is contradictory. Likewise, of 27 measures of students' social behavior, 17 support the hypothesis; again, not one finding favors large classes.

Psychological theory also explains why students may become better students in small classes. Four different perspectives support this hypothesis. The first might be called "the firing line hypothesis": in a small class, each student experiences continuing pressure to participate.

Any student may be called upon at any time to answer questions or participate in a class activity and so can't avoid the teacher's attention by sitting in an obscure place in the classroom. Moreover, the teacher can't readily ignore any particular pupil (even if he or she would like to do so).

Second, small classes tend to encourage a closer "sense of community" among students and between teachers and students.⁴² Teachers of small classes report that they know each individual student better than they would in a larger class. Students tend to be more supportive of one another and to de-



"We'd better be in for Ms. Resnick. You never know when you might need a good lifeline."

velop a stronger sense of identification with the class as a whole.

Third, the concepts of "social loafing" and "diffusion of responsibility" have been used to explain why individuals in smaller groups are often more responsive than individuals in larger groups. These concepts have been demonstrated for "critical incidents." For example, a stranger in need tends to get attention if there are one or two bystanders but tends not to get help from members of a larger crowd.⁴¹ The same principles may also apply to such settings as students in classrooms.

And fourth, the study of group dynamics has shown that the norms of small groups promote different behavior than do the norms of large groups. In fact, social psychologists have documented a negative correlation between the size of a group and its "functional size" — that is, the number of group members who participate in any given activity.⁴² Once again, these principles may apply to the classroom as well.

These propositions are only theory at this time. But all four perspectives lead to the conclusion that the intensity of the teaching/learning experience is increased for students in smaller classes. More research is needed to test these propositions. Between the empirical research base and psychological theory, there seems to be strong support for the hypothesis that the biggest change when class sizes are reduced is in student behavior.

A CONCLUDING COMMENT

Ongoing programs provide the opportunity for continued research on class size — work that is badly needed. In recent years, many districts have undertaken CSR initiatives, often without any accompanying research or evaluation. Given the extent of the existing research base, it may not be necessary to show that academic achievement is improved at every site, and further large-scale randomized experiments may be even less necessary.

However, it is necessary to make sure that smaller classes are implemented correctly and that problems are addressed quickly. The Buffalo Public Schools, with a significant CSR program, conducted an exemplary evaluation.⁴³ It focused on program implementation so that administrators could identify problems as they arose and make useful midcourse corrections. Evaluations can also provide the basic information needed by administrators, parents, and legislators to decide

whether resources have been allocated properly.

There is still much to learn about small classes and classroom processes. CSR sites provide researchers with a rare opportunity — a large number of "natural laboratories" for answering questions about implementation, processes, and outcomes. Educators working in settings where class sizes are reduced should encourage the administrators to engage in formative evaluation and research — for their benefit and for the benefit of the broader education community.

1. The issue of class size has been raised in many court cases in recent years. The January 2001 decision in *Campaign for Fiscal Equity (CFE), Inc. v. State of New York*, however, went further than all the others. Having found the state in violation of its constitution, which requires the provision of a "sound basic education," Justice Leland DeGrasse listed seven resources that must be provided to students to meet this requirement. First on the list is "sufficient numbers of qualified teachers, principals, and other personnel." Second is "appropriate class sizes."

2. Gene V. Glass and Mary L. Smith, *Meta-analysis of Research of the Relationship of Class Size and Achievement* (San Francisco: Far West Laboratory for Educational Research and Development, 1978); and Robert E. Slavin, "Achievement Effects of Substantial Reductions in Class Size," in idem, ed., *School and Classroom Organization* (Hillsdale, N.J.: Erlbaum, 1989), pp. 247-57.

3. Glen E. Robinson, "Synthesis of Research on the Effects of Class Size," *Educational Leadership*, April 1990, pp. 80-90.

4. Ibid.

5. See, for example, Alan B. Krueger, "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics*, vol. 114, 1999, pp. 497-532; Larry V. Hedges, Barbara Nye, and Spyros Konstantopoulos, "The Long-Term Effects of Small Classes: A Five-Year Follow-Up of the Tennessee Class Size Experiment," *Educational Evaluation and Policy Analysis*, vol. 21, 1999, pp. 127-42; and Harvey Goldstein and Peter Blatchford, "Class Size and Educational Achievement: A Review of Methodology with Particular Reference to Study Design," *British Educational Research Journal*, vol. 24, 1998, pp. 255-68.

6. Elizabeth Word et al., *Student/Teacher Achievement Ratio (STAR): Tennessee's K-3 Class-Size Study* (Nashville: Tennessee State Department of Education, 1990); Jeremy D. Finn, *Class Size and Students at Risk: What Is Known? What Is Next?* (Washington, D.C.: Office of Educational Research and Improvement, National Institute on the Education of At-Risk Students, U.S. Department of Education, 1998); and Jeremy D. Finn and Charles M. Achilles, "Tennessee's Class Size Study: Findings, Implications, Misconceptions," *Educational Evaluation and Policy Analysis*, vol. 21, 1999, pp. 97-110.

7. Jeremy D. Finn et al., "The Enduring Effects of Small Classes," *Teachers College Record*, vol. 103, 2001, pp. 145-83.

8. Alan B. Krueger and Diane M. Whitmore, "Would Smaller Classes Help Close the Black-White Achievement Gap?" Princeton University Industrial Relations Section Working Paper No. 451, Princeton, N.J., March 2001. This paper is available at www.irs.princeton.edu/wpframe.html.

9. The Tennessee state testing program for all students ends in grade 8, but there was no indication that the benefits would not continue into later grades.

10. There appears to be an anomaly in the grade-6 science data that may come from erroneous scoring or data entry. Close examination of the data did not reveal why it occurred.

11. Alex Molnar, Philip Smith, and John Zahorik, *1998-99 Evaluation*

Results of the Student Achievement Guarantee in Education (SAGE) Program (Milwaukee: School of Education, University of Wisconsin, December 1999), p. 3.

12. Alex Molnar et al., *1999-2000 Evaluation of the Student Achievement Guarantee in Education (SAGE) Program* (Milwaukee: School of Education, University of Wisconsin, December 2000).

13. For evaluations of the Burke County CSR initiative, see Pauline Egelson, Patrick Harman, and Charles M. Achilles, *Does Class Size Make a Difference? Recent Findings from State and District Initiatives* (Greensboro, N.C.: Southeast Regional Vision for Education [SERVE], 1996); and Pauline Egelson and Patrick Harman, "Ten Years of Small Class Size in Burke County, North Carolina," in Margaret C. Wang and Jeremy D. Finn, eds., *How Small Classes Help Teachers Do Their Best* (Philadelphia: Center for Research in Human Development and Education, Temple University, and U.S. Department of Education, 2000), pp. 227-78.

14. The consortium of organizations included the American Institutes for Research, the RAND Corporation, Policy Analysis for California Education, WestEd, and EdSource.

15. CSR Research Consortium, *Class Size Reduction in California: The 1998-99 Evaluation Findings* (Sacramento: California Department of Education, 2000).

16. Charles M. Achilles, Jeremy D. Finn, and Susan Gerber, "It's Time to Drop the Other Shoe: The Evidence on Teacher Aides," paper presented at the annual meeting of the Mid-South Educational Research Association, Bowling Green, Ky., November 2000; Jayne Boyd-Zaharias and Helen Pare-Bain, *Teacher Aides and Student Learning: Lessons from Project STAR* (Arlington, Va.: Educational Research Service, 1998); Jeremy D. Finn et al., "Teacher Aides: An Alternative to Small Classes?," in Wang and Finn, pp. 131-73; and Susan B. Gerber et al., "Teacher Aides and Students' Academic Achievement," *Educational Evaluation and Policy Analysis*, vol. 23, 2001, pp. 123-43.

17. Gerber et al., op. cit.

18. Finn et al., "Teacher Aides: An Alternative to Small Classes?"

19. Ibid., p. 163.

20. Ibid., pp. 163-64.

21. Ibid., p. 165.

22. Glass and Smith, op. cit.; and John M. Johnston, "What Are Teachers' Perceptions of Teaching in Different Classroom Contexts?," paper presented at the annual meeting of the American Educational Research Association, Boston, 1990.

23. Molnar, Smith, and Zahorik, op. cit.

24. CSR Research Consortium, op. cit.; and Charles M. Achilles et al., "A Study of Reduced Class Size in Primary Grades of a Fully Chapter 1-Eligible School: Success Starts Small (SSS)," paper presented at the annual meeting of the American Educational Research Association, San Francisco, 1995.

25. Jeremy D. Finn et al., "Carry-Over Effects of Small Classes," *Peabody Journal of Education*, vol. 67, 1989, pp. 75-84; and Carolyn M. Everison and John K. Folger, "Small Class, Large Class: What Do Teachers Do Differently?," paper presented at the annual meeting of the American Educational Research Association, San Francisco, 1989.

26. Barbara Harvey, "An Analysis of Grade Retention for Pupils in K-5" (Doctoral dissertation, University of North Carolina, Greensboro, 1993); and Word et al., op. cit.

27. Helen Pare-Bain, B. DeWayne Fulton, and Jayne Boyd-Zaharias, "Effects of Class Size Reduction in the Early Grades (K-3) on High School Performance: Preliminary Results (1999) from Project STAR," paper presented at the National Press Club, Washington, D.C., April 1999.

28. Krueger and Whitmore, op. cit.

29. Dominic Brewer et al., "Estimating the Cost of National Class Size

Reductions Under Different Policy Alternatives," *Educational Evaluation and Policy Analysis*, vol. 21, 1999, pp. 179-92.

30. Krueger, op. cit. Henry Levin conducted an independent analysis of these variables, presented at the 1998 meeting of the American Educational Research Association. Although the analysis was not published, Levin obtained figures very close to Krueger's.

31. Brian Stecher et al., "Class-Size Reduction in California: A Story of Hope, Promise, and Unintended Consequences," *Phi Delta Kappan*, May 2001, pp. 670-74.

32. Ibid.

33. Lorin W. Anderson, "Why Should Reduced Class Size Lead to Increased Student Achievement?," in Wang and Finn, pp. 3-24.

34. For elaboration on these points, see Gina M. Pannozzo and Jeremy D. Finn, "Professional Development and Support Needs of Class Size Reduction Teachers," paper presented at the National Invitational Conference on Taking Small Classes One Step Further, Washington, D.C., December 2000.

35. Personal communication.

36. Eugene M. Lewit and Linda S. Baker, "Class Size," *The Future of Children*, vol. 7, 1997, pp. 112-21; and Karen H. Miles, "Free Resources for Improving Schools: A Case Study of Teacher Allocation in Boston Public Schools," *Educational Evaluation and Policy Analysis*, vol. 17, 1995, pp. 476-93.

37. Jennifer McGiverin, David Gilman, and Chris Tillitski, "A Meta-Analysis of the Relation Between Class Size and Achievement," *Elementary School Journal*, vol. 90, 1989, pp. 47-56; and Daniel J. Maeler, Clinton I. Chase, and James J. Walden, "Effects of Reduced Class Size in Primary Classes," *Educational Leadership*, February 1988, pp. 48-53.

38. See Molnar et al., op. cit.

39. See, for example, Eric A. Hanushek, *The Evidence on Class Size* (Rochester, N.Y.: W. Allen Wallis Institute of Political Economy, University of Rochester, 1998).

40. Other economists have called into question Hanushek's conclusion that there is "no relationship" between resources and student achievement, showing that more appropriate analyses of his data—even based on pupil/teacher ratios—lead to the conclusion that resources are related to student achievement. See, for example, Larry V. Hedges, Richard D. Laine, and Rob Greenwald, "Money Does Matter Somewhere: A Reply to Hanushek," *Educational Researcher*, May 1994, pp. 9-10; and Alan B. Krueger, "An Economist's View of Class Size Research," in Wang and Finn, pp. 99-130.

41. Krueger, "An Economist's View"; and Krueger and Whitmore, op. cit.

42. Helen V. Bareman, "Students' Sense of Community: Implications for Classroom Size," paper presented at the National Invitational Conference on Taking Small Classes One Step Further, Washington, D.C., December 2000.

43. John M. Darley and Bibb Latane, "Bystander Intervention in Emergencies: Diffusion of Responsibility," *Journal of Personality and Social Psychology*, vol. 10, 1968, pp. 202-14; and John M. Levine and Richard L. Moreland, "Small Groups," in Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey, eds., *The Handbook of Social Psychology*, Vol. 2, 4th ed. (New York: McGraw-Hill, 1998), pp. 415-69.

44. Robert M. Bray, Norbert L. Kerr, and Robert S. Atkin, "Effects of Group Size, Problem Difficulty, and Sex on Group Performance and Member Reactions," *Journal of Personality and Social Psychology*, vol. 36, 1978, pp. 1224-40.

45. Jeremy D. Finn, Susan B. Gerber, and Gina M. Pannozzo, *Evaluation of the Class Size Reduction Initiative, Buffalo Public Schools, 1999-2000* (Buffalo: Graduate School of Education, State University of New York at Buffalo, August 2000).

The Enduring Effects of Small Classes

JEREMY D. FINN AND SUSAN B. GERBER

State University of New York at Buffalo

CHARLES M. ACHILLES

Eastern Michigan University

JAYNE BOYD-ZAHARIAS

HEROS, Inc.

The purpose of this investigation was to extend our knowledge of the effects of small classes in the primary grades on pupils' academic achievement. Three questions were addressed that have not been answered in previous research: (1) How large are the effects of small classes relative to the number of years students participate in those classes? (2) How much does any participation in small classes in K-3 affect performance in later grades when all classes are full-size? (3) How much does the duration of participation in small classes in K-3 affect the magnitude of the benefits in later grades (4, 6, and 8)? Rationales for expecting the continuing impacts of small classes were derived in the context of other educational interventions (for example, Head Start, Perry Preschool Project). The questions were answered using data from Tennessee's Project STAR, a statewide controlled experiment in which pupils were assigned at random to small classes, full-size classes, or classes with a full-time teaching assistant. Hierarchical linear models (HLMs) were employed because of the multilevel nature of the data; the magnitude of the small-class effect was expressed on several scales including "months of schooling." The results for question (1) indicate that both the year in which a student first enters a small class and the number of years (s)he participates in a small class are important mediators of the benefits gained. The results for questions (2) and (3) indicate that starting early and continuing in small classes for at least three years are necessary to assure long-term carryover effects. Few immediate effects of participation in a class with a full-time teacher aide, and no long-term benefits, were found. The results are discussed in terms of implications for class-size reduction initiatives and further research questions.

The issue of class size is near and dear to the heart of teachers, parents, and educational policy makers. None would argue that larger classes are better, and most would assert the obvious advantages of small classes over large.

Parents with adequate resources pay for their children to attend private schools, at least in part to reap the benefits of the small-class environment.

Recent decades have produced well over 100 empirical studies of class size. Because the studies used nonexperimental procedures, and because many involved small samples or were of short duration, few definitive conclusions could be drawn. Tentative conclusions were summarized in several widely read reviews, specifically the Glass-Smith meta-analysis (1978), and reviews by Educational Research Service (Robinson and Wittebols, 1986; Robinson, 1990) and Slavin (1989). The reviews converged on four major propositions: (1) "Reduced class size can be expected to produce increased academic achievement" (Glass and Smith, 1978, p. 4), although the effects of even substantial reductions are small (Slavin, 1989). (2) "The major benefits from reduced class size are obtained as the size is reduced below 20 pupils" (Glass and Smith, 1978, p. v). (3) Small classes are most beneficial in reading and mathematics in the early primary grades (Robinson, 1990). (4) "The research rather consistently finds that students who are economically disadvantaged or from some ethnic minorities perform better academically in smaller classes" (Robinson, 1990, p. 85).

In 1985, the Tennessee legislature funded an experiment—Project STAR (Student-Teacher Achievement Ratio)—to provide more definitive answers.¹ Project STAR was a controlled scientific experiment that built on the principles identified in prior research. Students entering kindergarten were assigned at random to a small class (S, 13–17 students), a regular class (R, 22–26 students), or a regular class with a full-time teacher aide (RA) within each participating school. The within-school randomization controlled for a host of between-school differences, including differences in the populations served, differences in per-pupil expenditures and instructional resources, and differences in the composition of the school staff. Teachers were assigned to the classrooms at random. The class arrangement was maintained throughout the day, all year long. There was no intervention other than class size and teacher aides.

Children were kept in the same experimental conditions (S, R, or RA) for up to four years, through grade 3. A new teacher was assigned to the class each year. Over 6,000 students in 329 classrooms in 79 schools in 46 districts participated in STAR in the first year, and almost 12,000 students participated in the course of the four-year study. An array of outcome measures was administered at the end of each school year, including both norm-referenced and criterion-referenced achievement tests in reading, mathematics, and other school subjects.

All pupils returned to full-size classes in grade 4. Fortunately, through funding for the Lasting Benefits Study (LBS) by the Tennessee State Department of Education and recent work by HEROS, Inc. (Pate-Bain et al., 1997), STAR participants have been followed through their high school years and

beyond. In addition to annual tests of academic achievement, student behavior was assessed in grades 4 and 8, attitudes toward school were assessed in grade 8, and school experiences were recorded throughout the grades.

Project STAR has already made major contributions to research on educational processes. It provided the most definitive answers to date about the effects of attending a small class in the primary grades. Many districts and states, the U.S. Department of Education, and several other countries have used STAR findings to guide class-size reduction initiatives. Further, STAR demonstrated that high-quality experimentation in education is feasible. It motivated decision makers to investigate ways to strengthen educational interventions by conducting further experiments on policy issues.² Finally, STAR produced a remarkable database now being used by research teams to answer further questions related to class size, questions about child development generally, and a host of other questions to guide educational policy.

The present investigation addressed three primary questions—questions not answered in previous work. First, we reexamined academic achievement for grades K through 3 to determine the extent to which the effects of small classes or teacher aides are related to the *number of years* that students participate in these settings. At the same time, we examined the impact of small classes on in-grade retentions during these years. The remaining questions pertained to long-term carryover effects of small classes: How much does any participation in small classes in K–3 affect performance in later grades, when children return to full-size classes? And how much does the *duration* of participation in small classes in K–3 determine the magnitude of continued benefits (in grades 4, 6, and 8)? Because samples and methods, not to mention the state of prior research, differ somewhat for short-term and long-term outcomes, the analyses are described in separate sections of this paper.

PART 1: IMMEDIATE EFFECTS OF SMALL CLASSES (GRADES K–3)

The original analyses of STAR data (Word et al., 1990; Finn and Achilles, 1990) consisted of cross-sectional analyses of achievement at the end of each year of experimentation, using nested ANOVA and MANOVA models. These analyses showed that students in small classes had superior academic achievement to students in regular-size classes in every school subject in every grade (K–3); high statistical significance was found for every test, subtest, and multivariate set of tests, including both the norm-referenced and criterion-referenced batteries.³ Further, in each grade, there were some significant interactions with urbanicity and/or race: Minority students and students attending inner-city schools reaped the greatest benefits of attend-

ing small classes. No differences were found between regular-size classes and classes with a full-time teacher aide.

Several factors suggest that the effects of small classes originally reported were understated. Because of student mobility and in-grade retentions, the STAR sample became successively more complex in grades 1, 2, and 3 (see STAR sample below). To the extent permitted by STAR data, both of these issues were addressed in the present study. Despite the complexities, all prior analyses of the K-3 results found significant benefits of small classes in all four years.

Concomitant studies of teaching practices and children's behavior in small classes revealed a set of mechanisms that help explain these differences. Students who attended small classes displayed improved learning behavior, increased engagement in school, and decreased disruptive or withdrawn behavior compared to their counterparts in regular-size classes (Finn, 1998; Finn, Fulton, Zaharias, and Nye, 1989). Teachers in STAR small classes spent increased time in direct instruction, and less time on managerial/organizational tasks (Evertson and Folger, 1989), a finding replicated in the year-long observational study, *Success Starts Small* (Achilles, Kiser-Kling, Owen, and Aust, 1994).

The present investigation reexamined the cross-sectional data, adding three features. First, we partitioned the sample more finely to examine the number of years a pupil participated in a small, regular, or teacher-aide class; this analysis also allowed us to control for student mobility during the experimental years. Second, we used hierarchical linear modeling (HLM) procedures to determine if the original findings replicate using an alternative statistical approach. Third, we focused on the question "How large is the effect?" by estimating strength-of-effect measures in terms of scale scores and in terms of months of schooling, a scale familiar to most educators.

METHODS

The STAR Sample

The selection of schools resulted in over 6,300 kindergarten students in 325 classrooms in 79 schools in 46 districts participating in the first year of the study (see Word et al., 1990; Finn and Achilles, 1990). Students and teachers were assigned at random to a small class (S), regular-size class (R), or regular-size class with a full-time teacher aide (RA) within each school. Larger schools had more than one of each class type so that all kindergarten pupils were allocated. With few exceptions, students were kept in the same class grouping throughout the years they participated in the experiment (for up to four years).

Kindergarten was not mandatory in Tennessee in 1985, so the number of students in the grade-1 sample was larger. Most (at least 90%) had attended some form of kindergarten, but not necessarily in a STAR school. Students entering STAR schools in grade 1, and others in grades 2 and 3, were also assigned to the three class types at random. Sample sizes and the composition of the original STAR sample are given in Table 1. Of the students identified as minority, 98.7% were African American.

Several factors caused the composition of the classes to become more complex in each successive grade. At the end of kindergarten, approximately one-half of regular-class students were assigned at random to teacher-aide classrooms and approximately one-half of teacher-aide students were assigned to regular classrooms.⁴ No students were purposely moved into or out of small classes, and no further reassignments were made after this point.

Second, migration of students into and out of STAR schools—a fact of life in both regular and experimental school programs—added to the complexity. Most grade-1 students had attended kindergarten but some had not. A small number of students, by changing schools, moved into a STAR classroom of a different type. The mixture of students in some classrooms became more complex in grades 2 and 3, with some students having attended the same type of class for zero, one, or two previous years. Krueger (1999) performed a careful analysis of student migration during the four-year experiment and concluded that it did not bias the class-size effect. Approximately 55% of kindergarten and grade-1 entrants participated in the experiment for three or more consecutive years.

Finally, in-grade retentions contributed to the complexity of the sample. Students retained in grade during the study were lost from STAR for sub-

Table 1. Sample sizes for experimental years (K-3)

	Grade			
	K	1	2	3
Number of schools	79	76	75	75
Number of classes	325	339	340	338
Small	127	124	133	140
Regular	99	115	100	90
Aide	99	100	107	108
Median class size				
Small	15	15	15	16
Regular	22	22	23	24
Aide	23	23	23	24
Number of students	6325	6829	6840	6802
Percent minority	32.8	33.1	34.6	33.4
Percent receiving free lunch	48.3	50.2	48.8	48.4

sequent years. While the number of retained students was not very great in total, fewer were retained in small classes compared to regular classes (Harvey, 1993). Interviews with teachers indicated that teachers in small classes had more confidence that students could be passed to the next grade and would receive the additional academic support they required. This left fewer spaces in small classes to accommodate retainees from the previous year. Project STAR did not collect the data needed to estimate the impact of student retentions fully; in the present study, we used the student's age to create a proxy for retention.

Sample sizes for our analyses, after eliminating students missing critical information, ranged from 5,394 to 5,910 for particular test batteries in particular grades.

Measures

The reading and mathematics scales of the Stanford Achievement Tests (SATs; Psychological Corporation, 1983) were administered in the spring of each year. This study used the scale scores created for SAT through item response theory (IRT) that can be compared across grades. For strength-of-effect measures, we also used the grade-equivalent (GE) scale. A grade equivalent of 2.4, for example, means that the pupil is performing like a typical student in the 4th month (December) of grade 2. Thus, if the student is actually in the 4th month of grade 1, he/she is performing quite well. If the student who took the test is actually in the 10th month (June) of grade 2, he/she is performing like students who have had 6 months less schooling. Grade equivalents can be obtained directly from the tables given in the publisher's manuals, or by fitting a curve of mean or median scale scores to the year and month of schooling in which the test was taken (see, for example, Shulz and Nicewander, 1997). The use of GEs has been subject to some debate, focused largely on the interpretation of individual students' scores (see Burket, 1984; Hoover, 1984; Peterson, Kolen, and Hoover, 1989; Yen, 1986). It is clear that GEs are not appropriate for estimating "rate of growth" because the scale is tied to the month/year metric; average growth for a cohort is always about 1.0 GE per school year. However, they are a useful way to compare the means of several groups at a particular grade level, and can be interpreted in terms familiar to educators—months of schooling.

Beginning in first grade, the Basic Skills First (BSF) tests, a set of curriculum-referenced tests developed by the State of Tennessee, were also administered to each student. These were constructed from well-specified lists of objectives in reading and mathematics at each grade level. A student was considered to have mastered an objective if (s)he answered 75% of the items correctly. The numbers of objectives in the tests were as follows—

grade 1: 8 reading and 11 mathematics; grade 2: 12 reading and 15 mathematics; grade 3: 10 reading and 15 mathematics. The present study used the number of objectives mastered in each subject in analyzing the BSF results.

Each pupil was coded according to the number of years spent in a particular class type ("duration"). For example, in grade 1, students in small, regular, or teacher-aide classes were coded to indicate whether this was their first or second year in that class type. In grade 2, duration was coded one, two, or three. Students with particular combined experiences were eliminated from the analysis. This includes students who, for whatever reason, had participated in both a small and teacher-aide classroom and students who had been in a small or teacher-aide classroom in previous years but were now in a regular class. As a result, the numbers of students omitted from the grade 1, 2, and 3 cross-sectional analyses were 230, 477, and 662, respectively.⁵

In each grade, the student's age was used to indicate the likelihood that (s)he had been retained in a previous grade. Students who were 14 months or more above the grade-appropriate age were classified as "over age." For example, Tennessee students were admitted to first grade if they reached their sixth birthday (72 months) by September 1. Any first grade student 86 months or older on that date was placed in the over-age category. The limited retention data gathered in STAR indicated that most over-age students had been retained in a previous grade, a finding confirmed by Harvey (1993) from a sample of actual student records. The student's participation in the free lunch program was used as an indicator of socioeconomic status (SES).

Analyses

A three-level HLM analysis was performed for each of the five achievement tests separately at each grade level. Level-1 (student) variables were gender (male-female), race (white-minority), the numerical duration index, the interaction of race with duration, and the test score. Level-2 (classroom) variables included two class-type contrasts: small classes compared to regular classes and teacher-aide classes compared to regular classes. Level-3 variables were two dummy codes to compare three school locations: suburban and rural schools compared to inner-city schools, respectively. The interactions of class type with race, duration, and urbanicity were included in the model as well as the interactions of race with duration, urbanicity with duration, and the three-way race-by-duration-by-class type and urbanicity-by-duration-by-class type interactions.⁶ All tests were conducted at the .01 significance level. When the three-way interactions and some two-way interactions were found to be nonsignificant, the models were reestimated retaining only the significant effects.

Several follow-up analyses were performed. First, SES was added to the final model and all analyses were rerun to see if there were any noteworthy changes in the results; the SES contrast was coded as "high-low" (full-price lunch-free lunch). Second, specific contrasts were tested in the grade 1, 2, and 3 data not represented by individual regression coefficients. In grade 1, we compared the performance of students who had been in small classes (and teacher aide classes) for one year or two consecutive years with the performance of grade-1 regular-class students. In grade 2, we compared the performance of students who had been in small classes (and teacher-aide classes) for one year, two years, or three years, respectively, with the performance of students who had been in regular classes; and so on.⁷ One, two, and three years in a small class or teacher-aide class were viewed as different treatment conditions to be compared to the "norm" (students in regular classes). As a final set of analyses, the over-age variable was added to the reduced model to see if achievement was affected above and beyond the effects of different retention rates.

Several aspects of the analysis addressed the issue of student mobility. The duration variable, obtained for both regular and experimental students, provided some statistical control because students with shorter durations in STAR were likely to have transferred from another school. Further, the analysis included three indicators linked closely with mobility: student race, SES, and school urbanicity.

Effect-size measures were derived from the final models as follows. Predicted means ("fitted means") for small and regular classes were obtained from the reduced HLM model. For both SAT and BSF tests, the mean difference was divided by the standard deviation of all students in regular classes in that grade. Next, the predicted means on the SAT tests were converted to grade equivalents (GEs) using a curve derived from the tables published by the Psychological Corporation. These were subtracted to estimate the difference in months of schooling.

IMMEDIATE EFFECTS: RESULTS

The final HLM regression coefficients for grades K-3, including the control for SES, are given in Appendix Table A-1. Differences on background characteristics were consistent with prior research on the performance of elementary-grade pupils. White students performed significantly better than minorities on all achievement tests in all grades, and higher-SES students performed better than lower-SES students. In fact, the SES effect was substantially greater than the racial difference in most instances. Girls performed significantly better than boys on all verbal measures (Total Reading, Word Study Skills, BSF Reading) in each grade. Although girls performed better than did boys in mathematics in kindergarten, the differences were

not significant in grade 1 and were significant only on the BSF Mathematics test in grades 2 and 3. All comparisons of school locations were statistically significant. Students in suburban and rural settings performed better than did students in inner-city schools on every test in each grade.

Overall Differences among Class Types

Overall differences among class types (labeled "Small" in Table A-1) were consistent. In every grade (K-3), students attending small classes performed better academically—on all achievement tests—compared to students in regular-size classes.⁸ In grade 3, when tests were administered in other subject areas, the same results were found on the Language, Science, and Social Science tests as well. No overall differences were found between teacher-aide classes and regular classes (labeled "Aide" in Table A-1) at the .01 level of significance.

The HLM analyses also revealed several significant or marginally significant interactions with class type. Interactions in grade 1 and 2 involved the duration variable discussed below. In grade 3, the urbanicity-by-small interaction was significant for Word Study Skills. The means⁹ indicate that this occurred because inner-city students benefitted more from small classes than did students in other locations, achieving a small-class advantage nearly twice as large as that of students in suburban and rural schools.

How large were the benefits of small classes? To answer this question, effect sizes were estimated for the small class-regular class difference (Table 2). The magnitude of the overall small class advantage was about 0.2σ on each test in kindergarten, increasing to about 0.3σ in grade 1 and then declining slightly in grades 2 and 3. The decreased effect sizes in grades 2 and 3 reflect in part that these samples include pupils who attended small classes for different numbers of years—from 1 to 3 years in grade 2 and from 1 to 4 years in grade 3.

The grade-equivalent results for the SAT tests estimate the benefits of attending a small class in "months of schooling." At the end of kindergarten, students in small classes were about one-half month ahead of their peers in reading and about 1.6 months ahead in mathematics. Each of these increased in grade 1. Grade-1 students who attended small classes were about 1.3 months ahead of regular-class students in reading and about 2.8 months ahead in mathematics. In grade 2, small-class students were between 3.5 and 4.8 months ahead of students in regular classes. In grade 3, the advantage was still greater in verbal areas but appeared to drop somewhat in mathematics.¹⁰ In general, the advantage of small classes was both statistically significant in every grade and, in months of schooling, increased from year to year.

Table 2. Effect sizes for small classes

Grade	SAT tests ^a			BSF tests ^a		SAT grade equivalents ^b		
	Total Reading	Word Study Skills	Total Mathematics	Reading	Mathematics	Total Reading	Word Study Skills	Total Mathematics
Kindergarten	0.21	0.20	0.19	—	—	0.5	0.5	1.6
Grade 1	0.30	0.29	0.31	0.31	0.21	1.3	0.9	2.8
Grade 2	0.26	0.25	0.25	0.21	0.18	3.9	4.8	3.5
Grade 3	0.22	0.21	0.15	0.14	0.11	4.5	5.5	2.6

^aDifference between mean performance of small-class students and regular-class students, divided by standard deviation of all regular-class students.

^bDifference between small-class students and regular-class students in months of schooling.

Differences According to the Number of Years in Small Classes

Students in grades 1, 2, and 3 were classified according to the number of years they had attended a particular class type. The small class-regular class comparisons are displayed in Figure 1; detailed HLM results for the analysis of number of years are given in Table A-1. In Figure 1, the vertical axis corresponds to the effect size, that is, the number of standard deviations separating the mean performance of students in small classes and the mean of students in full-size classes. The dots are placed at a height corresponding to this axis. For the SAT tests, each dot corresponding to a statistically significant difference is also labeled with the effect size in terms of months of schooling (GE).

The upper-left portion of Figure 1 shows results for the Reading scale of the Stanford Achievement Tests. In kindergarten, the advantage of attending a small class was about 0.21σ or one-half month on the grade equivalent scale; this difference is significant at $p < .001$. In grade 1, students attending small classes for the first time performed better in reading than students in regular classes. The effect size was 0.16σ , corresponding to an advantage of about 6/10 of a month; the difference was significant at $p < .01$. Grade-1 students who had attended small classes for two years (since kindergarten) had a greater advantage over regular-class students. The effect size was 0.40σ , corresponding to an advantage of about 1.9 months of schooling; this difference was significant at $p < .001$.

In grade 2, the reading effect sizes for students who attended small classes for one, two, and three years were 0.12σ , 0.24σ , and 0.36σ , or advantages of approximately 1.8 months, 3.7 months, and 5.7 months, respectively. These differences were significant at the .05 and .001 levels. In grade 3, the advantage was not statistically significant for students who attended small classes for the first time, but was statistically significant for those who had attended for one or more previous years. The advantages of attending a small class for two years (since grade 2), for three years (since grade 1), and for four years (since kindergarten) were approximately 3.3 months, 5.2 months, and 7.1 months, respectively.

The same pattern was obtained for every achievement test (including Word Study Skills; see Table A-1): In each grade, there was a clear benefit of attending small classes for additional years. Grade-1 students who were in small classes for the first time scored between 0.6 and 2.0 months ahead of regular-class students on the SAT tests; those who attended small classes for two years scored between 1.9 and 3.4 months ahead of regular-class students. In grade 2, the advantage of having spent three years in a small class was substantially greater than two years, which was, in turn, greater than the advantage of being in a small class for the first time. In grade 3, the advantage of having been in a small class for four years reached 7.1

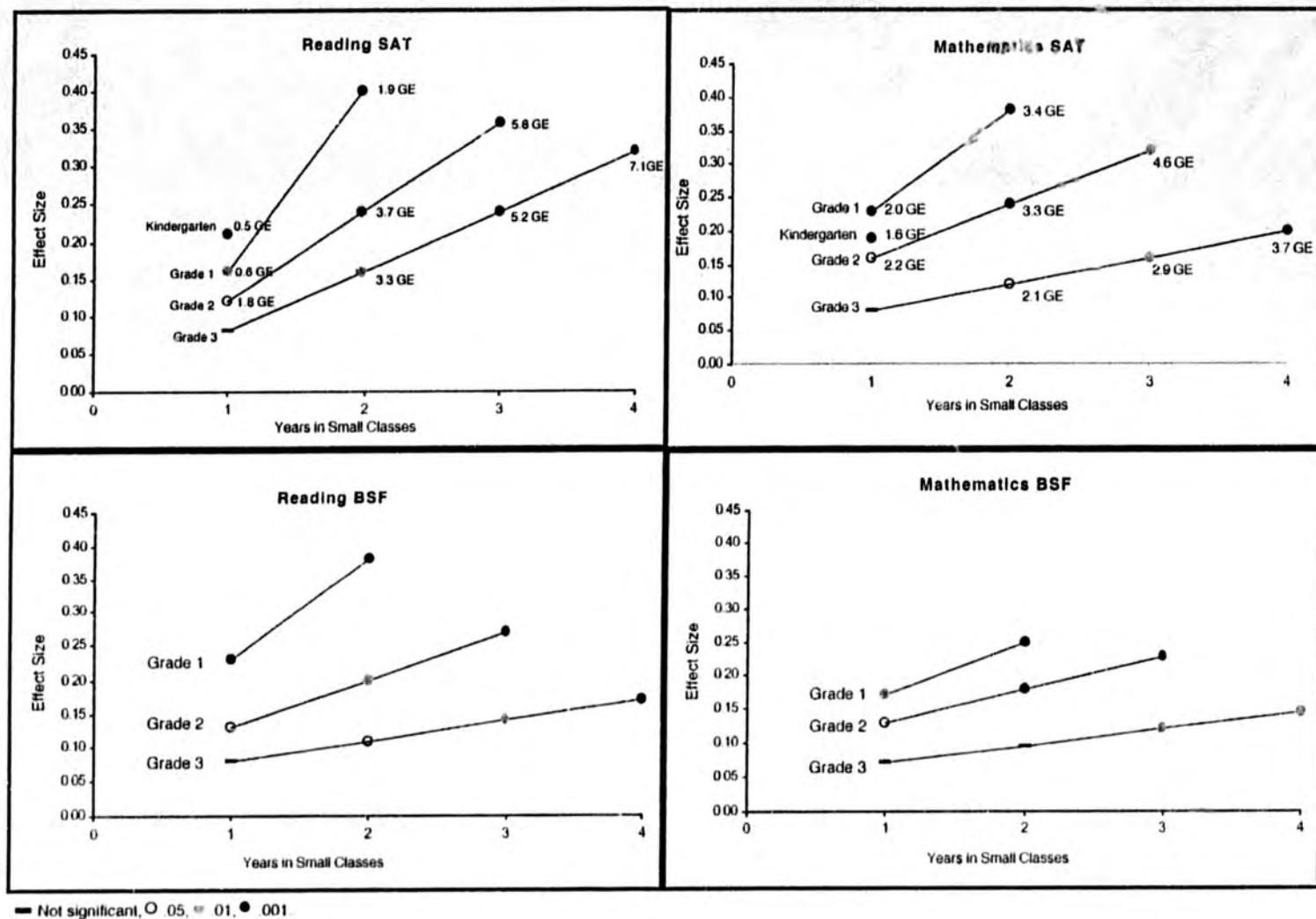


Figure 1: Small-class effects by number of years of participation (duration)

months and 8.2 months in the verbal areas and 3.7 months in mathematics. Effect sizes for the BSF tests also increased monotonically as students spent additional years in a small class.¹¹

The effect sizes shown in Figure 1 and the significance levels in Table A-1 demonstrate two principles: The first is the importance of continuous attendance in small classes. It is clear from the figure that the effects increase with each additional year in a small class. In addition, in later grades, the effects were only statistically significant for those students who had continued small-class participation. In grade 2, none of the one-year comparisons was statistically significant at $p < .01$ (although several were "marginally" significant). In contrast, students who attended small classes for two years (grades 1 and 2) or for three years (grades K, 1, and 2) performed significantly better than did students in regular classes. Likewise, grade-3 students attending small classes for the first time performed better, but not significantly so, compared with students in regular classes. Those who attended small classes for two years (grades 2 and 3) performed significantly better than regular-class students on two of the tests (SAT Reading and Word Study Skills). Students who attended small classes for three or four years performed significantly better than regular class students on all achievement tests.

The second principle is the importance of starting small classes in the earliest grades. Note, for example, that one year in a small class was significantly related to students' performance in kindergarten and grade 1, less significant in grade 2, and not significant in grade 3. Likewise, two years in a small class had significant impacts on performance in grade 1 and grade 2 but produced mixed results in grade 3. Thus, the earlier small classes are introduced, the greater the potential for a strong impact on academic achievement.

Other findings. Several specific interactions with duration were also statistically significant; we examined the interaction means for each. In grade 1, the duration-by-race-by-small interaction occurred because minority students who attended small classes for two years scored substantially closer to white students who attended small classes for two years on the BSF reading test. This reduced achievement gap was not found for any other combination of race, duration, or class type. In grade 2, the duration-by-urbanicity-by-small interaction occurred because inner-city students who attended small classes for three years scored closer to suburban and rural students who attended small classes for three years on the BSF reading test. Again the achievement gap was reduced for students in inner-city schools.

The overall comparison between teacher-aide classes and regular classes was not significant for any test in any grade. However, several specific

contrasts were statistically significant. In grade 1, students who had teacher aides for two years outperformed their regular-class counterparts on the verbal tests (SAT Reading and Word Study Skills; BSF Reading). In grade 2, students who had teacher aides for three years outperformed their regular-class counterparts on the SAT verbal tests. In grade 3, no teacher-aide effects attained the .01 significance level. We view these results as worthy of note and further exploration. It is possible, for example, that aides are most useful in the reading areas during the years when reading is emphasized (particularly in grade 1).

In-Grade Retentions

Project STAR did not collect extensive information about student retentions; we used the "over-age" variable as a proxy. In kindergarten, 3.2% of students in small classes were over-age in comparison to 2.6% of students in regular and teacher-aide classes. Although the random assignment to the three conditions was carried out and monitored carefully, slightly more retainees or late entrants were assigned to small classes. In grade 1, 7.5% of small-class students were over-age in comparison to 13.1% of regular students and 13.7% of teacher-aide students. These pupils may have been retained either in kindergarten or in grade 1 the previous year. In grade 2, the percentages were 12.4% in small classes compared to 19.2% and 20.1% in regular and teacher-aide classes. In grade 3, the percentages were 16.9% in small classes compared to 21.1% and 21.5% in regular and teacher-aide classes.

Because fewer students were retained in small classes than in regular or aide classes, there was less space for retained students and fewer retainees were assigned to small classes each year. Differential retentions are both an outcome of small class participation and a factor that may confound the analysis of academic achievement. Thus, we repeated the final analyses of achievement (including SES) with over-age as an additional control variable.

The difference in performance between over-age and non-over-age students was highly significant ($p < .001$) on every test in every grade, with over-age students having substantially poorer achievement scores. The tests of significance among class types, however, showed no consistent or noteworthy change from the uncontrolled analysis. Some p -values were unchanged—for example, all tests of significance for small classes in grades K and 1. Others increased—for example, several tests of particular contrasts in grades 2 and 3. And others decreased slightly—for example, the three- and four-year comparisons for the two mathematics tests in grade 3.

Likewise, the teacher aide findings were not changed in the reanalysis, although several contrasts showed less statistical significance. In grade 1, the two-year contrast for BSF Reading did not reach the .01 significance

level. In grade 2, all of the three-year contrasts for SAT tests became non-significant; and in grade 3, the largest four-year contrast (SAT Reading) was reduced substantially.

In general, when we asked whether small classes had an impact on academic achievement above and beyond the effects of in-grade retention, the answer was "yes." Controlling for retentions, a number of comparisons reached higher levels of significance. In addition to confirming the pattern of overall effectiveness of small classes, the analysis suggested that even one year in a small class in grade 2 or 3 is worthwhile for many students. On the other hand, the apparent effectiveness of teacher aides for students in aide classes for three or four years may be a function of differential retentions rather than a true benefit to pupil performance.

CONCLUSIONS

The reexamination of data from Tennessee's Project STAR for grades K to 3 substantiate three conclusions. First, on average, students in small classes perform better than do students in regular classes or regular classes with teacher aides in each grade on all tests of academic performance. The "typical" magnitude of the small-class advantage was approximately 0.2σ on every test for the total sample of pupils in each grade. This finding parallels earlier work using other statistical approaches (for example, Finn and Achilles, 1990), although we did not find as many significant interactions with race/ethnicity or urbanicity as found previously. Those interactions that were significant or approached statistical significance were all in the direction of greater small-class advantages for students who usually have lower performance, that is, minorities or students attending inner-city schools.

Second, both the year in which a student first enters a small class and the number of years (s)he participates in a small class are important mediators of the benefits gained. For students who began small classes in kindergarten or grade 1, there was a significant advantage in both reading and mathematics. For those who attended small classes in *both* kindergarten and grade 1, the small class advantage was greater still, putting these students from approximately 1.3 months to 3.4 months ahead of their counterparts who attended full-size classes.

In grades 2 and 3, the benefits of attending a small class were not statistically significant unless students had attended for at least two years. Again, in each grade, the benefits of additional years in a small class were greater. In grade 2, for example, the achievement advantage for students who had participated in small classes for one, two, and three years was 0.12σ , 0.24σ , and 0.36σ , respectively, in reading and 0.16σ , 0.24σ , and 0.32σ , respectively, in mathematics. In terms of months of schooling, these advantages placed the students from 1.8 months, to 3.7 months, to 5.8

months ahead of their peers in reading, and from 2.2 months, to 3.3 months, to 4.6 months ahead in mathematics. The results were not diminished when statistical control was introduced for the number of over-age students enrolled in each class.

Third, in general, we found few if any academic benefits associated with a full-time teacher aide. Several statistically significant differences were found for grade-1 students who had attended teacher-aide classes for two years, and several were found for grade-2 students who had attended for three years. The significance levels were reduced, in some cases to nonsignificance, when statistical control for student retentions was added to the analysis. More research on the specific duties of teacher aides and how they may interact with retained or low-performing students may be useful.

PART 2: ENDURING EFFECTS OF SMALL CLASSES (GRADES 4-8)

Project STAR ended after grade 3, and all students returned to regular classes in grade 4. Achievement and behavior data continued to be collected, allowing us to ask how strong the benefits of small classes would be after pupils return to full-size classes. We began our study of carryover effects with a question of theory: Under what conditions would we *expect* the benefits to endure into later grades?

Other early-childhood programs have had mixed results. It is not uncommon that immediate cognitive benefits, reflected in tests of academic achievement, tend to decrease over time (White, 1985-86). There is some evidence of lasting benefits for non-achievement outcomes, however (Barnett, 1995; Lazar and Darlington, 1982). Both the Perry Preschool Project (Berrueta-Clement et al., 1984) and most Head Start programs exhibit this pattern.¹² In both of these programs, experimental and comparison-group students were indistinguishable on tests of academic achievement three years after students left the program. However, there were indications that program students were less likely to be placed in special education, less likely to be retained in grade, and more likely to graduate than their comparison-group counterparts (Berrueta-Clement et al., 1984; McKey et al., 1985).

The Chicago Parent Child Centers (CPC) program documented continuing effects on academic achievement (Reynolds, 1997). Designed to aid low-income students, especially those not served by Head Start, the programs have components for preschool, kindergarten, and grades 1 through 3. Thus, children can participate for as many as six years. Preschool programs meet for one-half day, but other portions are full-day programs. Class activities, focusing on the development of language and reading skills, take place in small classes. Each classroom also has a full-time teacher aide, and

parental involvement is extensive. An evaluation of continuing effects showed that CPC students outperformed non-program students in reading and mathematics in grades 3 and 5, and in mathematics in grade 8; effect sizes ranged from .17 σ to .34 σ (Reynolds, 1997). Like those in the Perry Preschool Project and Head Start, students attending the CPC were less likely to be retained in grade.

What features of educational programs are likely to foster long-term benefits? Summarizing the evaluations of 36 early childhood programs, Barnett (1995) concluded that to have any long-term effects at all, school-age services "must actually change the learning environment in some significant ways" (p. 43). Ramey and Ramey (1998) forwarded a conceptual framework for early interventions and identified six principles of efficacy. The principles rest on the concept that "fragmented, weak efforts in early intervention are not likely to succeed, whereas intensive, high-quality, ecologically pervasive interventions can and do" (p. 109). The primary principles are: (1) developmental timing, that is, start early and continue; (2) program intensity, that is, the importance of many hours per day, days per week, and weeks per year of the intervention; and (3) direct provision of learning experiences, rather than relying on intermediary sources; parent training alone, for example, is not likely to have an enduring impact on children's learning.

The principle of developmental timing is consistent with all that is known about the importance of intervening when the rate of growth is greatest, and the increasing difficulty of effecting change as children grow older (for example, Bloom, 1964). Studies of program duration support the principle of program intensity. In one, Head Start-only students were compared with others who remained for an additional four years in the continuation, Project Follow Through (Abelson, Zigler, and DeBlais, 1974). Students in the extended program outperformed their counterparts on several aspects of the Peabody achievement tests, leading the authors to conclude that "gains accruing from compensatory education programs are commensurate with duration" (p. 770). An evaluation of CPC programs (Reynolds, 1997) examined both starting age and duration. Students who participated in the CPC for one or two years were indistinguishable from the control group on grade-8 achievement tests, while those who participated for five or six years were superior in both mathematics and reading. The advantage of duration remained significant controlling for age of entry, while age of entry was not significant controlling for duration. Duration in the intense CPC program was more important than age of entry.

The Perry Preschool Project and most Head Start programs are of limited intensity (see Zigler and Styfco, 1994). Although beginning at an early age (Perry at age 3, Head Start at 4), the Perry program lasted for two years,

ending when pupils entered regular schools at age 5. Likewise, most Head Start programs begin when the child is 4 and end when regular school begins in kindergarten. Neither program engages pupils for the full day. The Perry intervention involved about $2\frac{1}{2}$ hours of school time daily; typical Head Start programs involve about $3\frac{1}{2}$ hours of class time, four or five days per week. Both extend the hours of engagement somewhat through home visits, although there is much variability in this practice among Head Start sites (Zigler and Styfco, 1994).

When students leave programs such as Perry, CPC, or Head Start, they are likely to enter half-day kindergartens with programs targeted to average pupils. The advantages of full-day kindergartens—especially those with a heavy academic emphasis—are well documented (Achilles, Nye, and Bain, 1994–1995; Cryan *et al.*, 1992; Finn, n.d.; Karweit, 1989; Naron, 1981). Yet by 1997, only 10 states¹³ required all districts to offer full-day programs (Council of Chief State School Officers, 1998). The remaining states required half-day programs (21), required half-day or full-day programs at the discretion of individual districts (10), or had *no* statewide policy about the provision of kindergarten (10).

Although directed to all students—not just children of poverty—Tennessee's Project STAR also started early. Beginning with full-day kindergartens, STAR was a high-intensity intervention, affecting children for the entire school day every day of the school year, for up to four consecutive years. STAR impacted the learning setting directly and thus influenced all student-teacher interactions taking place in that setting. We would expect that (1) there would be positive impacts of small classes on academic achievement in the later grades, and (2) the long-term impact would be related to the duration of a student's participation in small classes in the early grades.

METHODS

Samples and Measures for Enduring Effects

Beginning in the fall of 1989, when STAR pupils were in grade 4, Tennessee initiated the Tennessee Comprehensive Assessment Program (TCAP). Every pupil was tested on the TCAP battery, which included norm-referenced tests from the Comprehensive Tests of Basic Skills (CTBS/McGraw Hill, 1989) and BSF tests for each grade in reading and mathematics. In the first year of testing, some districts declined to participate, resulting in a smaller follow-up sample for grade 4. However, most or all districts participated in the TCAP program by the time STAR pupils reached grades 6 and 8.

The present study used the IRT scale scores from the CTBS Total Mathematics, Total Reading, Science, and Social Science scales, and the number

of objectives mastered on the BSF reading and mathematics tests as outcome measures. The total numbers of objectives on the test were 8 and 8 in grade 4, 7 and 9 in grade 6, and 10 and 8 in grade 8 for reading and mathematics, respectively.

The full sample for the study of enduring effects included students who had attended STAR regular classes or else small or teacher-aide classes for one or more years. A few students who attended both a small and teacher-aide class at some point (for example, by changing schools) were eliminated. The grade-4 sample comprised 4,015 students in 61 schools providing data on the CTBS and 4,045 students providing data on the BSF tests (see Table 3). While these schools had participated in STAR the previous year, only 40% of inner-city schools provided scores for the grade-4 analysis. The grade-6 and grade-8 samples were larger and more dispersed as students entered middle- and high-school grades. The original proportions of inner-city, suburban, and rural schools were reestablished. The grade-6 sample comprised 6,100 students in 518 schools who provided CTBS results and 2,737 students in 213 schools who provided BSF scores.¹⁴ The grade-8 sample comprised 5,835 students in 489 schools who provided CTBS results and 5,217 students with BSF scores. Although the number of schools is large, most pupils were concentrated in a subset of them. For example, in

Table 3. Sample sizes for post-experimental years

	Grade		
	4	6 ^a	8
Number of schools	61	518	489
Number of students (CTBS tests/BSF tests)	4015/4045	6100/2737	5835/5217
Percent minority	19.8	32.1	32.4
Percent receiving free lunch	39.3	47.1	47.0
Number of schools in duration sample	61	464	405
Number of students in duration sample (CTBS tests)			
Small classes, by duration			
1 year	72	277	263
2 years	66	174	182
3 years	181	250	223
4 years	637	693	657
Regular classes	506	923	899
Teacher aide classes	1371	1989	1940

^aBSF data only available for 2737 students in 213 schools in Grade 6 in full sample. For duration analysis, Grade-6 BSF data only available for 2017 students in 158 schools.

sixth grade, 229 schools had five or more STAR pupils; these schools accounted for 90.3% of our grade-6 sample. In eighth grade, 183 schools had five or more STAR pupils, accounting for 90.9% of our sample.

To study duration (Research Question 3), a sub-sample of students was selected who entered Project STAR in kindergarten or grade 1 and who participated in the same class type (small, regular, or teacher aide) for one year or for two, three, or four consecutive years. Kindergarten and grade-1 entrants constituted the original experimental plan and had the opportunity to participate for up to four years. Students who entered STAR classes in grade 2 or 3 and those with other patterns of small-class participation were eliminated from this analysis;¹⁵ the resulting samples were smaller, due primarily to the elimination of students who entered STAR in grade 2 or 3 (see Table 3).¹⁶

Analyses

The full-sample analysis examined overall differences among students who had been in small, regular, or teacher-aide classes in K-3. This consisted of testing a series of HLM models for two levels of data, students and schools. Student-level variables included gender (male-female), race (white-minority), two class-type contrasts (small-regular; aide-regular), and the test score. School-level variables included two contrasts among three school locations: suburban and rural schools compared to inner-city schools, respectively. The interactions of class type with race and with urbanicity were included in the model as well. When interactions were found to be non-significant, the "reduced" main-effects model was rerun and SES (full-price lunch-free lunch) was added to see if any noteworthy changes in the results occurred. Effect sizes were estimated in standard deviation units and in grade equivalents, using the final models in the same manner as in the K-3 analysis.

To study the effects of duration, we classified each pupil according to the grade of entry into STAR (K or 1) and the extent of continuous participation. Students were identified who had one year of participation in a small or teacher-aide class (K or 1), two years in sequence (K-1 or 1-2), three years in sequence (K-1-2 or 1-2-3), or four years. Students who never participated in either experimental condition were classified as regular-class pupils and were also classified by the number of years they had been in Project STAR.

The study of duration was complicated by the fact that students who left STAR before reaching third grade, due to their family moving or to being retained in grade, were not a representative cross-section of all pupils. For example, students who move from school to school are generally from

lower-income families; that, and changing schools itself, may result in reduced academic achievement. In all, 32.8% of students who entered STAR kindergarten classes were minority and 48.3% participated in the free-lunch program. Of those who left STAR classes before reaching third grade, 39.9% were minority and 58.6% were receiving free lunches. Similarly, 39.3% of students who entered STAR classes in grade 1 were minority and 61.2% were receiving free lunches. Of those who left STAR classes before third grade, 43.6% were minority and 63.8% were receiving free lunches. Statistical control for school urbanicity and student race, SES, and duration of participation in STAR helped tighten the internal validity of the conclusions.

Preliminary analyses indicated a confounding of grade of entry with the effects of duration, for several reasons. Students who entered in grade 1 could not participate in the experiment for more than three years. Also, when students were cross-classified by starting grade, duration, and class type, the grade-4 sample had a number of very small cell *ns*. Every analysis, regardless of test, grade, or other variables in our models, indicated that both short-term and long-term effects were mediated by the number of years a student participated in a small class. For these reasons, buttressed by the CPC findings that duration was more important than starting age (Reynolds, 1997), we completed our analyses by focusing on duration of small-class participation.

The primary analyses consisted of testing HLM models with two levels of data, students and schools. The only school-level variable was urbanicity. Student-level variables included the outcome measure plus gender (male-female), race (white-minority), SES (full-price lunch-free lunch), class type (small, regular, or aide), and the numerical duration index (1, 2, 3, or 4). The full model also included interactions with class type and duration (class type-by-duration; class type-by-race; class type-by-urbanicity; duration-by-race; duration-by-urbanicity; duration-by-class type-by-race; duration-by-class type-by-urbanicity). In each grade, contrasts were tested that were not represented by the individual regression coefficients. Specifically, we compared the performance of students who had been in small classes (and teacher-aide classes) for one year, two years, three years, and four years, respectively, with the performance of students who had been in regular classes.¹⁷ Again, one, two, three, and four years in a small class or teacher-aide class were viewed as different treatment conditions to be compared to the "norm" (regular classes). Students with shorter durations in Project STAR may have transferred to other schools and thus comprised a selectively mobile group. The inclusion of race, SES, urbanicity, and duration in Project STAR in the statistical models provided as much control for mobility as possible from our data set.

As in the other analyses, when some or all interactions were found to be nonsignificant, they were eliminated and effect size measures were derived from the reduced models. Predicted means were expressed as effect sizes using the standard deviations of all regular-class students, and these were converted to grade-equivalents using curves derived from the tables published by CTBS/McGraw Hill.

ENDURING EFFECTS: RESULTS

How Much Does Any Participation in Small Classes in K-3 Affect Performance in Later Grades?

The reduced-model HLM coefficients for grades 4, 6, and 8, including the control for SES, are given in Appendix Table A-2. Results for student and school demographics are consistent with the K-3 analysis. Race and SES differences were highly statistically significant on virtually all achievement tests in all three grades. Girls maintained their superiority over boys in reading and mathematics, with very few exceptions, through grade 8, and performed significantly better than did boys in social sciences in grade 8. Boys exhibited superior performance in science in grades 4 and 8.¹⁸ Students in suburban and rural schools performed significantly better than did their inner-city counterparts on every test in all three grades.

Overall differences among class types. Grade-4 students who had attended small classes for one or more years in K-3 scored significantly better on all tests of academic achievement than did students who attended full-size classes; all differences were significant at $p < .01$ or below despite the smaller grade-4 sample size. In grade 6, all differences were significant in favor of small classes except for science. In grade 8, differences favoring small classes were significant at $p < .01$ or below on four of the six achievement tests, and approached significance ($p < .05$) on the other two. No significant effects were found in grade 4, 6, or 8 for students who attended teacher-aide classes.

The small class-regular class differences are given in Table 4. In grade 4, the small-class advantage ranged from 0.11σ to 0.15σ on the CTBS tests and 0.14σ and 0.16σ on the BSF tests. In grade 6, apart from science, the small-class advantages ranged from 0.10σ to 0.20σ , being generally about the same or slightly less than the values for grade 4. In grade 8, the small-class effect sizes ranged from 0.08σ to 0.14σ . In general, effect sizes remained in about the same range from grade 4 to grade 8.

In terms of months of schooling (grade equivalents), the advantage of having attended a small class for one or more years ranged from 2.4 months

Table 4. Overall small-class effect sizes for grades 4, 6, 8

Grade	CTBS tests ^a				BSF tests		CTBS grade equivalents ^b			
	Total Reading	Total Mathematics	Science	Social Science	Reading	Mathematics	Total Reading	Total Mathematics	Science	Social Science
Grade 4	0.15	0.13	0.11	0.12	0.14	0.16	3.9	2.4	3.1	4.6
Grade 6	0.10	0.10	0.02 ^c	0.12	0.20	0.13	3.0	3.1	1.0	5.1
Grade 8	0.11	0.08 ^d	0.09	0.08 ^d	0.14	0.09	4.1	3.4	4.3	4.8

^aDifference between mean performance of small-class students and regular-class students, divided by standard deviation of all regular-class students.

^bDifference between small-class students and regular-class students in months of schooling.

^cNot significant.

^dMarginally significant ($p < .05$).

to 4.6 months in grade 4, from 3.0 to 5.1 months in grade 6 (apart from science), and from 3.4 months to 4.8 months in grade 8. In general, students who attended small classes during the K-3 period were at least 2.5 months ahead in all school subjects, and as much as five months ahead in some, compared to their counterparts who attended regular classes.

Attending a small class in grades K-3 is associated with enduring academic benefits in all school subjects in grades 4, 6, and 8. It is likely that the impact endures beyond grade 8; unfortunately, we did not have achievement test data for higher grades. The question "How large is the effect?" is answered partially from these results. The effect size measures represent *lower limits* on expected outcomes because the sample included students who entered STAR in different grades and participated in small classes for as little as one year and others who participated for two, three, or four years. Despite the diverse nature of the sample, virtually all results were statistically significant and the total lasting impact of small classes, having affected all school subjects for five years or more, was substantial.

How Much Does the Duration of Participation in Small Classes in K-3 Affect Benefits in Later Grades?

The duration sub-sample consisted of those participants who entered Project STAR in kindergarten or grade 1 and who could be classified as having one year of participation or two, three, or four contiguous years. Although the duration sub-sample was about 30% smaller than the full samples in grades 4, 6, and 8, tests of significance for background characteristics (race/ethnicity, SES, gender, and urbanicity) produced results similar to those for the full samples; they are not tabled again. And like the full-sample analysis, the duration analysis produced no teacher aide-regular class contrast that was significant at $p < .01$ in any grade.

Differences according to the number of years in small classes. Figure 2 is a display of the differences in performance between students who attended small classes and those who attended full-size classes; the organization of Figure 2 is the same as that of Figure 1.

The upper-left portion of Figure 2 shows results for the Total Reading scale of the CTBS. In grade 4, the difference between the average performance of students who had been in small classes for one year and students who had been in regular classes was about 0.04σ (not statistically significant). The effect size for students who had been in small classes for two years was approximately 0.12σ (not statistically significant). The effect size for three years in a small class was approximately 0.20σ . This difference corresponds to an advantage of approximately 4.6 months of schooling

(GE) (statistically significant at the .01 level). Four years in a small class produced an advantage of 0.28σ or approximately 6.6 months of schooling (significant at the .001 level).

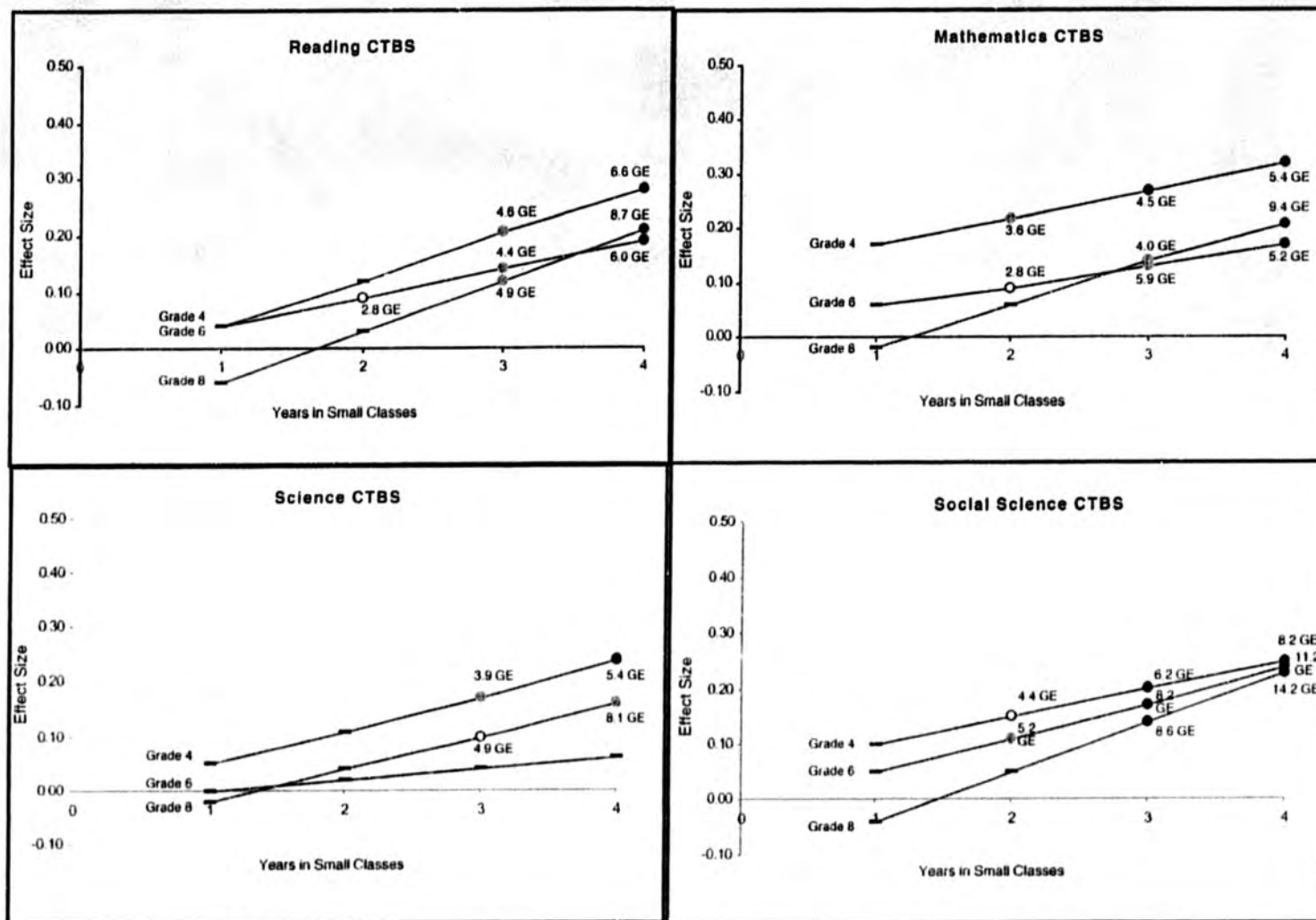
The patterns in grade 6 and grade 8 are similar. Effect sizes for one year in a small class are small and nonsignificant. Effect sizes for two years in a small class are larger but either marginally significant or nonsignificant. Effect sizes for three years in a small class are 0.14σ in grade 6 and 0.12σ in grade 8, corresponding to a 4.4-month advantage and a 4.9-month advantage in reading, respectively; both are statistically significant at the .01 level. The greatest effects were found for four years in a small class, which peaks at approximately a 6.0-month advantage in grade 6 and an 8.7-month advantage in grade 8.

The same pattern was obtained for every CTBS achievement scale and the criterion-referenced BSF tests. Specifically, students who had attended small classes for one year performed no better in grade 4, 6, or 8, on average, than students who had attended regular-size classes. Although there were significant "immediate" benefits of participating in a small class in kindergarten or grade 1 alone, the benefits did not endure through the later grades when all students returned to full-size classes. The one-year effect sizes are small and some are even slightly negative; none is significant at $p < .01$.

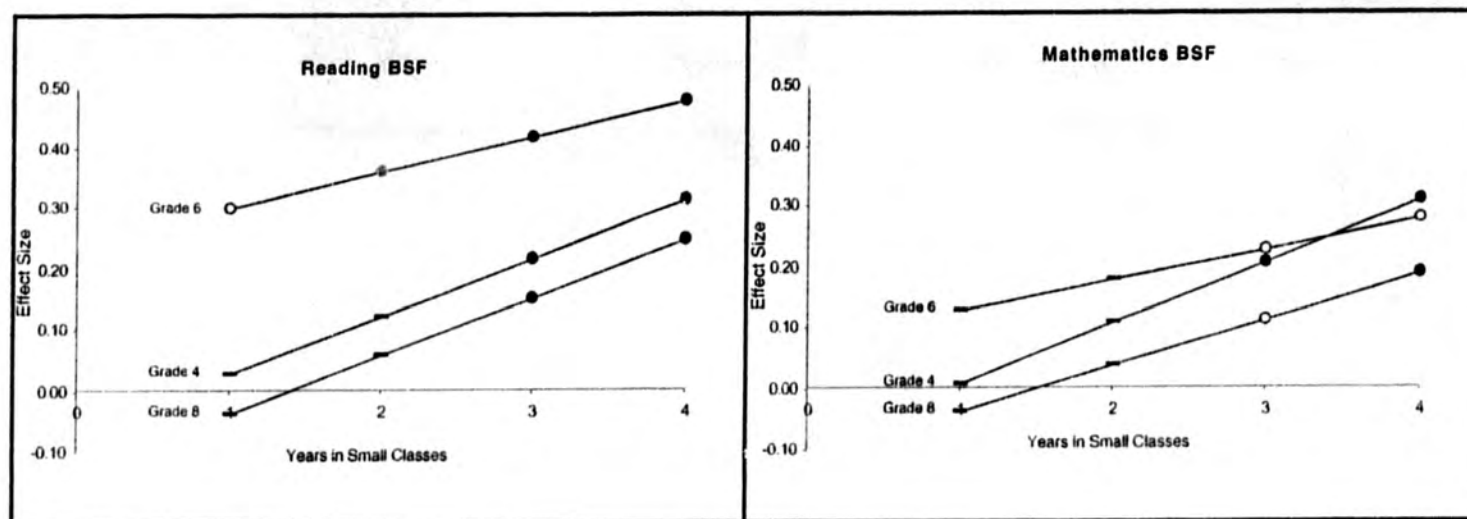
Two years in a small class (K-1 or 1-2) yielded consistently larger effect sizes than did one year, but only attained statistical significance for some tests in some grades (CTBS Mathematics in grade 4 and CTBS Social Science and BSF Reading in grade 6). Although there may have been some carryover effects of two years in a small class, they were minimal by the time two, four, or six additional years had gone by. Effect sizes for two years in a small class are all larger than for one year in a small class, however; the median effect size is $.10\sigma$ considering grades 4, 6, and 8 together.

Three years in a small class (K-1-2 or 1-2-3) had important carryover effects in later grades. In grade 4, differences favoring small classes were statistically significant on all CTBS and BSF scales. In both grades 6 and 8, differences favored small classes on all tests except CTBS Science and BSF Mathematics. Every effect size for three years in a small class is larger than the respective effect size for two years in a small class. The median effect size is 0.21σ in grade 4, 0.15σ in grade 6, and 0.13σ in grade 8, considering all six tests. On the grade-equivalent scale, the median advantage for students who spent three years in a small class is about 4.5 months in grade 4, about 4.2 months in grade 6, and about 5.4 months of schooling—over one-half of a school year—in grade 8.

The carryover effects of four years in a small class (K-1-2-3) were highly statistically significant on all tests in all later grades, with the single exception of grade 6 science. Every effect size for four years in a small class was



— Not significant, ○ .05, ⊙ .01, ● .001.



— Not significant, ○ .05, □ .01, ● .001.

Figure 2: Small-class carryover effects by number of years of participation

larger than the respective effect of three years in a small class. The median effect size was 0.29σ in grade 4, 0.21σ in grade 6, and 0.21σ in grade 8. On the grade-equivalent scale, the median advantage for students who spent four years in a small class was about 5.4 months in grade 4, 5.6 months in grade 6, and 9.0 months in grade 8. In grade 8, students are almost a full school year ahead of their classmates who attended larger classes in K-3.

In contrast to the lower limits estimated from the entire grades 4, 6, and 8 samples, these effect sizes represent *best estimates* of impact on students who attend small classes for one, two, three, or four consecutive years. While one year in a small class did not have a consistently significant impact on later school performance, two years in a small class had somewhat more, and three and four years showed lasting benefits that are statistically and educationally meaningful. Improvements in test scores remained significant through grade 8—fully five years after the small classes were disbanded. Few educational interventions have demonstrated this degree of longevity.

SUMMARY AND CONCLUSIONS

Prior to Project STAR, research on class size was limited largely to nonexperimental studies, most of which were based on small samples or of short duration. To the credit of Tennessee legislators and with the cooperation of schools across the state, a remarkable four-year experiment was designed and executed (see Ritter and Boruch, 1999, for a discussion of the political origins of STAR). Because of its design and magnitude, the results of Project STAR eclipsed most of the work that preceded it.

The original STAR reports demonstrated that small classes in the primary grades have a positive impact on academic achievement in all subject areas in kindergarten through third grade (for example, Word et al., 1990; Finn and Achilles, 1990). Perhaps of equal importance, the project yielded an excellent database for addressing a plethora of additional questions. The present study examined one set of questions in depth: To what extent are the impacts related to the number of years a student attends a small class? We examined the effects on academic achievement both while students attended small classes and in later grades, after they returned to full-size classes.

STAR was a randomized experiment that manipulated and controlled only class size (small or regular) and one pupil-teacher ratio variable (regular and regular with an aide). Natural variations in schooling made STAR classes more heterogeneous each year after kindergarten. Student mobility into and out of STAR schools and in-grade retentions resulted in each class being comprised of students who had participated in a small, regular, or

teacher-aide classroom for different numbers of years. Although these processes occur commonly in elementary and secondary schools, they have the potential to bias statistical outcomes. Students who change schools more frequently are often from lower socioeconomic strata and suffer academically. Fewer students were retained in STAR small classes, leaving fewer "slots" for the previous year's retainees.

To study duration in small classes, we focused on just those students who spent one year, or else two, three, or four consecutive years, in a small, regular, or teacher-aide class. In examining the data during the experimental years (K-3), we used student age as a proxy for in-grade retention and asked whether benefits of small classes remained even after retentions were controlled statistically. For both the experimental years and post-experimental years, we addressed student mobility as best the data would allow by controlling for factors related to mobility, that is, student race, SES, and school urbanicity.

The immediate benefits of small classes were clear and consistent. Students who attended small classes performed significantly better on all achievement measures in all grades than did students in full-size classes with or without teacher aides. The number of years a student spent in a small class was also a determinant of the benefits realized.¹⁹ In each grade, having participated in a small class for more years produced greater gains than did fewer years. In grades 2 and 3, in fact, the advantages were only statistically significant when a student had been in small classes for two or more years. Those students were generally 3 to 8 months ahead of their peers who attended large classes, depending on the specific test and grade level.

Starting early is also important. One year of participation in a small class had a strong effect on pupil performance in kindergarten and grade 1, but less impact in grades 2 and 3. Two years in a small class had a profound impact on achievement in grade 2 but less effect in grade 3. In general, the later the starting point, the more years of participation are required to improve students' learning. Starting point and duration cannot be separated clearly in the STAR data, however. Attending small classes for two or more years by grade 2, for example, implies that the student began in a small class by kindergarten or first grade. Thus we are left with the conclusion that entering a small class in the early grades (K or 1) and continuing for two or more years has a meaningful, significant impact on students' academic achievement.

Writing about non-school settings, social psychologists have documented that individual participation is inversely related to group size; that is, the more people present, the less "involved" each individual tends to be (for example, Darley and Latane, 1968; Levine and Moreland, 1998). STAR-related research suggests that this principle may help explain the benefits

of small classes as well. Pupils in small classes were found to invest more effort toward learning activities, take greater initiative in the classroom, and display less disruptive or inattentive/withdrawn behavior than their counterparts in full-size classes (Finn, 1998; Finn et al., 1989). It appears that the very feature of "smallness" sets the stage for increased student engagement in learning. At the same time, when the class is smaller, teachers are compelled to attend to all students regardless of their abilities, motivation, or classroom demeanor.

In addition to immediate impact, attending small classes also had long-term benefits. In general, students who attended small classes in K-3 performed better academically on all subjects in grades 4, 6, and 8 than their peers who attended full-size classes (see also Nye, Hedges, and Konstantopoulos, 1999). Our expanded analysis of starting point and duration in a small class revealed other considerations. One year in a small class in kindergarten or grade 1 was not sufficient to produce long-term effects, even through grade 4. Two years was somewhat better. Carryover effects were consistently significant only for students who had attended small classes for three to four years. Four years in a small class put students nearly a whole school year ahead of their counterparts who had attended larger classes in K-3. In STAR, four years in a small class was also confounded with starting age, since four-year participants in STAR must have entered small, regular, or teacher-aide classes in kindergarten.

We conclude that entering a small class in kindergarten or grade 1 and remaining in that setting for at least three years produces, on average, significant and noteworthy improvements in academic achievement at least through grade 8 in all school subjects. These findings are consistent with the criteria for lasting effects outlined by Ramey and Ramey (1998): The students started early, continued through a number of years, and their exposure to the program was maximal. Our findings underscore the importance of continued participation in small classes. They refute some detractors of small class policies who argue that one year, at most, produces all the benefits that can be realized from small classes (e.g., Hanushek, 1998). One year (in kindergarten or grade 1) produces early gains but does not produce lasting effects.

This investigation, for the most part, confirmed that attending a class with a full-time teaching assistant had little impact on academic performance in grades K-3 and no significant effects on performance in later grades. The benefits of paraprofessionals may lie in other domains, for example, helping a small number of students who have learning or behavior problems, performing tasks so that a teacher has more time to devote to instruction, or even serving as a culturally relevant link to parents and the community; these possibilities require further study. However, it is clear that adding a teaching assistant to a full-size class does not affect academic

achievement in general and is not an effective alternative to reducing class size if increased learning is the goal.

The results for teacher-aide classes also address a distinction that has sometimes been muddled. In discussing class size research, some authors have failed to distinguish between *pupil-teacher ratios* and the number of pupils in a particular classroom (for example, Hanushek, 1998; Hedges, Laine, and Greenwald, 1994; Wenglinsky, 1997). Pupil-teacher ratios, usually computed for entire school districts or states, include special education and Title I teachers with small classes, subject matter specialists with no full-time classes of their own, librarians, teaching assistants, and others. Aide classes in this study actually had smaller *pupil-teacher ratios* than did small classes: about 24-to-2 or 12-to-1 for aide classes and about 15-to-1 for small classes. Our study shows that small classes have academic benefits not obtained by reducing the pupil-teacher ratio—even in a single classroom.

Many questions remain. Some can be addressed through further analysis of the STAR and follow-up database. For example, Glass and Smith (1978) proposed that a class size "threshold" (fewer than 20 pupils) is required to obtain academic benefits. Natural variation in class sizes within the ranges defined as small or regular in Project STAR could be analyzed to answer the question "How small must a class be to produce the benefits that have been observed?" Also, the database could be used to examine school-to-school variability in the magnitude of the small-class effect. The benefits of small classes were greater in some schools, less in others, and even nonexistent in a few (see Krueger, 1999). Random regression models can be used to ask how much of this variability is due to such factors as school size, the composition of the student body, or the preparation of teachers.

Other questions require new research. For example, like most prior research on class size, Project STAR focused on the primary grades; less is known about the possible benefits of small classes in the middle grades or in high school, although Boozer and Rouse (1995) demonstrated some small-class effects in grades 7 and 8. Even more important, however, is the question of teaching to a smaller group: What can teachers do to take advantage of the opportunities a small class provides, to maximize the impact on pupil learning and behavior? Answers to this question are virtually nonexistent at present. However, well over half of the states and many districts across the country are currently undertaking class-size reduction initiatives; they span the primary, middle, and secondary grades. We urge educators and researchers to make use of these *natural laboratories* to study questions such as these. There is still much to be learned.

APPENDICES

Table A-1. HLM reduced-model results for experimental years (regression coefficients; *p*-values)

Grade/Effect	SAT tests			BSF tests	
	Total Reading	Word Study Skills	Total Mathematics	Reading	Mathematics
Kindergarten					
Gender	-4.9	-5.0	-5.9		
Race	7.1	6.7	17.2		
SES	14.7	15.8	19.8		
Class type					
Small	6.4	7.2	9.1	N/A	
Aide	1.1	1.3	0.2		
Urbanicity					
Suburban	13.5	16.9	20.6		
Rural	10.5	11.7	17.7		
Grade 1					
Gender	-10.1	-7.9	-0.2	-0.4	-0.0
Race	7.4	5.4	16.8	0.3	0.6
SES	26.9	24.5	17.6	0.7	0.6
Duration	9.3	7.6	4.8	0.3	0.1
Class type					
Small	16.0	15.8	13.0	0.6	0.5
1 year—regular	8.8	10.7	9.6	0.4	0.4
2 years—regular	21.6	19.8	15.8	0.7	0.5
Aide—Regular	5.6	5.5	2.2	0.2	0.1
1 year—regular	-2.0	0.1	-1.5	0.0	0.0
2 years—regular	11.5	9.7	5.0	0.3	0.2

Urbanicity					
Suburban	36.2	30.0	22.5	0.7	0.9
Rural	41.9	30.7	29.0	0.7	0.9
Interactions ^a				^b	

Grade 2

Gender	-8.4	-5.4	-0.2	-0.6	-0.3
Race	8.2	6.1	13.8	0.5	0.4
SES	21.4	22.7	18.6	1.0	0.8
Duration	5.2	6.7	2.4	0.2	0.1
Class type					
Small	11.5	12.5	10.9	0.6	0.5
1 year—regular	5.5	4.1	6.9	0.4	0.3
2 years—regular	10.9	11.6	10.5	0.6	0.4
3 years—regular	16.2	19.2	14.1	0.8	0.5
Aide	3.5	3.7	2.2	0.1	0.1
1 year—regular	-2.7	-5.0	-2.1	-0.1	0.0
2 years—regular	2.8	2.8	1.7	0.1	0.1
3 years—regular	8.3	10.5	5.6	0.3	0.2
Urbanicity					
Suburban	28.5	28.3	18.8	1.5	0.6
Rural	36.7	33.6	30.9	1.7	1.0
Interactions ^a				^c	

Grade 3

Gender	-7.8	-8.1	-0.8	-0.5	-0.2
Race	8.0	6.1	8.8	0.4	0.6
SES	15.6	15.3	14.9	0.9	1.1
Duration	3.0	3.4	2.5	0.1	0.1
Class type					
Small	8.1	9.2	6.0	0.3	0.4
1 year—regular	2.9	4.2	3.1	0.2	0.2
2 years—regular	6.0	7.2	4.8	0.3	0.3

continued

Table A-1. *continued*

Grade/Effect	SAT tests			BSF tests	
	Total Reading	Word Study Skills	Total Mathematics	Reading	Mathematics
3 years—regular	9.2	10.2	6.6	0.4	0.4
4 years—regular	12.4	13.2	8.4	0.4	0.5
Aide	0.7	1.0	-0.7	-0.0	-0.1
1 year—regular	-4.6	-3.8	-3.4	-0.1	-0.2
2 years—regular	-1.4	-0.9	-1.8	-0.1	-0.2
3 years—regular	1.8	2.0	-0.2	0.0	-0.1
4 years—regular	5.1	4.9	1.4	0.1	-0.0
Urbanicity					
Suburban	19.9	23.9	17.2	0.8	1.2
Rural	24.7	28.5	22.9	1.1	1.8
Interactions ^a		^d			

□ NS □ $p < .05$ □ $p < .01$ ■ $p < .001$

^aDuration-by-class type interaction also included in all models; required to estimate the contrasts of 1, 2, 3, and 4 years in a small-class setting or teacher aide setting with regular-class students. Not tabled since it does not directly test the interaction of other contrasts in the table.

^bMarginal duration-by-race-by-small interaction ($p < .02$).

^cSignificant race-by-aide and duration-by-urbanicity-by-small interactions ($p < .01$).

^dMarginal urbanicity-by-small interaction ($p < .02$).

Table A-2. HLM results for post-experimental years, full sample (regression coefficients; *p*-values)

Grade/Effect	CTBS tests			BSF tests		
	Total Reading	Total Mathematics	Science	Social Science	Reading	Mathematics
Grade 4						
Gender	-6.8	-8.4	5.1	-1.0	-0.6	-0.2
Race	12.5	7.5	12.1	7.7	0.7	0.7
SES	23.3	19.6	21.6	27.3	1.1	1.0
Class type						
Small—Regular	7.6	5.6	6.0	7.1	0.3	0.4
Aide—Regular	2.0	1.0	1.2	2.2	0.1	0.1
Urbanicity						
Suburban	32.7	17.7	38.2	33.3	1.7	1.3
Rural	34.0	22.6	42.4	37.9	1.7	1.5
Grade 6						
Gender	-0.4	-6.4	1.0	-1.2	-0.6	-0.3
Race	13.4	8.6	18.4	5.9	0.7	0.5
SES	19.9	20.8	21.9	19.8	1.2	1.1
Class type						
Small—Regular	4.2	4.6	1.3	5.1	0.5	0.3
Aide—Regular	-0.7	1.5	-2.1	1.9	0.2	0.1
Urbanicity						
Suburban	19.3	15.0	29.7	13.4	1.4	0.8
Rural	17.8	14.9	31.2	12.7	1.6	1.1
Grade 8						
Gender	-1.7	-5.6	8.5	-3.8	-0.7	-0.2
Race	16.0	12.0	15.8	12.9	0.9	0.7
SES	20.8	21.3	17.1	18.6	1.2	1.2
Class type						
Small—Regular	4.8	3.9	4.1	3.8	0.3	0.3
Aide—Regular	0.5	0.1	0.9	0.9	0.1	0.0
Urbanicity						
Suburban	25.7	21.7	26.7	27.7	1.5	1.6
Rural	27.7	24.3	28.0	28.7	1.6	1.7

□ NS □ $p < .05$ □ $p < .01$ □ $p < .001$

Portions of this paper were presented at the conference on the Economics of School Reform at the Hebrew University of Jerusalem, May 23–26, 1999. This work was supported by a grant from the Spencer Foundation entitled "A Study of Class Size and At-Risk Students." We are grateful to HEROS, Inc., for assistance in obtaining the data used in this investigation, and to Steven Raudenbush and John Willett for advice on the statistical methods.

Notes

1. Project STAR was directed by Ms. Elizabeth Word of the Tennessee Department of Education and conducted by a consortium of researchers from four Tennessee universities. The Principal Investigators were C. M. Achilles (University of Tennessee), H. P. Bain (Tennes-

see State University), J. Folger (Vanderbilt University), J. Johnston (University of Memphis, for three years), and Fred Bellott (University of Memphis, one year). Jeremy Finn was external evaluator for the duration of the project.

2 Several national conferences have been held on this topic in the United States, using STAR as the model.

3 The same findings were confirmed by Goldstein and Blatchford (1998) and by Krueger (1999) using other statistical approaches.

4 This was done to assuage the concerns of some parents whose children had neither small classes nor a teacher aide. By the end of four years, this reassignment was judged to be relatively unimportant because no differences were found between regular and teacher-aide classes.

5 These were already subtracted from the sample sizes reported above (that is, 5,394 to 5,910).

6 Student variables were centered at the classroom means, class variables at the school means, and school variables were grand-mean centered. All terms in the model were treated as fixed except for the student-, classroom-, and school-level intercepts, which were treated as random.

7 These were obtained by evaluating algebraically the "expected values" of the final models for particular values of the independent variables (for example, one year of small classes; two years of small classes; regular classes) and subtracting (see Bryk and Kaudenbush, 1992). This process yields a set of coefficients that multiply the regression weights to estimate each contrast of interest. The coefficients can be entered back into the HLM program—the easiest way to compute estimates and tests of the desired contrasts.

8 The small-regular contrast for BSF mathematics in grade 3 was not significant at the .01 level; the *p*-value was .017.

9 Not tabled.

10 This may represent a real decline in the advantage of small classes in mathematics, because the effect size also decreases in grade 3. Although the effect is still substantial, this apparent inconsistency should be examined through further research including an analysis of the mathematics curriculum in grades 1 through 3.

11 The criterion-referenced BSF tests do not lend themselves to grade-equivalent scores.

12 Note, however, that Head Start programs are varied and some have different patterns of outcomes from others.

13 Including the District of Columbia.

14 Due to technical errors, portions of the grade-6 BSF data were lost when the original files were created and are irretrievable. The students for whom data are available, however, comprise a substantial group who attended all three class types in all three school locations (inner-city, suburban, rural). We judged that including these data in the analyses would be worthwhile.

15 Grade-2 and grade-3 entrants also presented problems because of the absence of information about their prior schooling and class placements.

16 Grade-2 and grade-3 entrants comprised about 29% of the overall sample. Samples for regular classes were reduced at a higher rate due to the reassignment of students between teacher-aide and regular classes from kindergarten to grade 1; we classified pupils who participated in teacher-aide classes in any grade as "teacher-aide" students, to keep the definitions consistent.

17 The procedures used were the same as in the K-3 analysis.

18 We examined the data to try to explain the anomalous lack of gender differences on the grade-6 CTBS tests but could find nothing out of the ordinary in the distributions.

19 The median number of students in STAR "regular" classes was 22-24, fewer than the number in many classrooms in schools today. Krueger (1999) found a statistically significant negative correlation between actual class size and achievement among the STAR regular

classes (the control group). It is possible that reducing class sizes in many of today's schools would have larger effects than those found in STAR.

References

- Abelson, W. D., Zigler, E., & DeBlaise, L. L. (1974). Effects of a four-year Follow Through program on economically disadvantaged children. *Journal of Educational Psychology*, 66, 756-771.
- Achilles, C. M., Kiser-Kling, K., Owen, J., & Aust, A. (1994). *Success starts small: Life in a small class. Small Grant-Field-Based Research Final Report*. Greensboro, NC: University of North Carolina.
- Achilles, C. M., Nye, B. A., & Bain, H. P. (1994-95). The test-score "value" of kindergarten for pupils in three class conditions, at grades 1, 2, and 3. *National Forum of Educational Administration and Supervision Journal*, 12, 3-15.
- Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *The Future of Children*, 5 (3), 25-50.
- Berrueta-Clement, J. R., Schweinhart, L. J., Barnett, W. S., Epstein, A. S., & Weikart, D. P. (1984). *Changed lives: The effects of the Perry Preschool Program on youths through age 19*. Ypsilanti, MI: High Scope Press.
- Bloom, B. S. (1964). *Stability and change in human characteristics*. New York: John Wiley & Sons.
- Boozer, M., & Rouse, C. (1995, June). *Intraschool variation in class size: Patterns and implication* (Working paper no. 344). Washington, DC: National Bureau of Economic Research, Industrial Relations Section. ERIC document no. ED 385-935.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, NJ: Sage Publications.
- Burket, G. R. (1984). Response to Hoover. *Educational Measurement: Issues & Practice*, 3 (4), 15-16.
- Council of Chief State School Officers. (1996, October). *Key state education policies on K-12 education*. Washington, DC: Author.
- CTBS/McGraw Hill. (1989). *CTBS: Comprehensive tests of basic skills*. Monterey, CA: Author.
- Cryan, J. R., Sheehan, R., Wiechel, J., & Bandy-Hedden, I. G. (1992). Success outcomes of full-day kindergarten: More positive behavior and increased achievement in the years after. *Early Childhood Research Quarterly*, 7, 187-203.
- Darley, J. M., & Latane, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, 10, 202-214.
- Everton, C. M., & Folger, J. K., (1989, March). *Small class, large class: What do teachers do differently?* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Finn, J. D. (n.d.). *Full-day kindergarten: Answers with questions*. Philadelphia: Laboratory for Student Success, Temple University Center for Research in Human Development and Education.
- Finn, J. D. (1998). *Class size and students at risk: What is known? What is next?* Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Institute on the Education of At-Risk Students.
- Finn, J. D., & Achilles, C. M. (1990). Answers and questions about class size: A statewide experiment. *American Educational Research Journal*, 27, 557-577.
- Finn, J. D., Fulton, B. D., Zaharias, J., & Nye, B. A. (1989). Carry-over effects of small classes. *Peabody Journal of Education*, 67, 75-84.
- Glass, G. V., & Smith, M. L. (1978). *Meta-analysis of research of the relationship of class size and achievement*. San Francisco: Far West Laboratory for Educational Research and Development.
- Goldstein, H., & Blatchford, P. (1998). Class size and educational achievement: A review of methodology with particular reference to study design. *British Educational Research Journal*, 24, 255-268.

- Hanushek, E. A. (1998). *The evidence on class size*. Rochester, NY: University of Rochester, W. Allen Wallis Institute of Political Economy.
- Harvey, B. (1993, December). *An analysis of grade retention for pupils in K-3*. Unpublished doctoral dissertation. University of North Carolina, Greensboro.
- Hedges, L. V., Laine, R. D., & Greenwald, R. (1994). Money does matter somewhere: A reply to Hanushek. *Educational Researcher*, 23 (4), 9-10.
- Hoover, H. D. (1984). The most appropriate scores for measuring educational development in the elementary school: GE's. *Educational Measurement: Issues & Practice*, 3 (4), 8-14.
- Karweit, N. (1989). Effective kindergarten programs and practice for students at risk. In R. E. Slavin, N. L. Karweit, & N. A. Madden (Eds.), *Effective programs for students at risk* (pp. 103-142). Boston: Allyn & Bacon.
- Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics*, 114, 497-532.
- Lazar, I., & Darlington, R. (1982). *Lasting effects of early education: A report from the consortium for longitudinal studies*. Chicago: University of Chicago Press for the Society for Research in Child Development.
- Levine, J. M., & Moreland, R. L. (1998). Small groups. In D. T. Gilbert, S. T. Fiske, and G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 2, 4th ed., pp. 415-469). New York: McGraw Hill.
- McKey, R. H., Condelli, L., Granson, H., Barnett, B., McConkey, C., & Plantz, M. (1985, June). *The impact of Head Start on children, families, and communities*. Washington, DC: Head Start Bureau, U.S. Department of Health and Human Services. Available through the Government Printing Office.
- Naron, N. K. (1981). The need for full-day kindergarten. *Educational Leadership*, 38, 306-309.
- Nye, B., Hedges, L. V., & Konstantopoulos, S. (1999). The long-term effects of small classes: A five-year follow-up of the Tennessee class size experiment. *Educational Evaluation and Policy Analysis*, 21, 127-142.
- Pate-Bain, H., Boyd-Zaharias, J., Cain, V. A., Word, E. R., & Binkley, M. E. (1997, September). *The student/teacher achievement ratio project: STAR follow-up studies 1996-1997*. Lebanon, TN: HEROS, Inc.
- Peterson, N. S., Kolen, M. J., & Hoover, H. D. (1989). Scaling, norming, and equating. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 221-264). New York: Macmillan.
- Psychological Corporation, Harcourt Brace Jovanovich. (1983). *Stanford Achievement Test* (7th ed.) San Diego: Author.
- Ramey, C. T., & Ramey, S. (1998). Early intervention and early experience. *American Psychologist*, 53, 109-120.
- Reynolds, A. J. (1997). *The Chicago child-parent centers: A longitudinal study of extended early childhood intervention* (Discussion paper no. 1126-97). Madison, WI: Institute for Research on Poverty.
- Ritter, G. & Boruch, R. (1999). The political and institutional origins of the Tennessee experiment. *Educational Evaluation and Policy Analysis*, 21, 111-125.
- Robinson, G. (1990). Synthesis of research on the effects of class size. *Educational Leadership*, 47 (7), 80-90.
- Robinson, G. E., & Wittebols, J. H. (1986). *Class size research: A related cluster analysis for decision making*. Arlington, VA: Educational Research Service.
- Schulz, E. M., & Nicewander, W. A. (1997). Grade equivalent and IRT representations of growth. *Journal of Educational Measurement*, 34, 315-331.
- Slavin, R. E. (1989). Achievement effects of substantial reductions in class size. In R. E. Slavin (Ed.), *School and classroom organization* (pp. 247-257). Hillsdale, NJ: Erlbaum.
- Wenglinsky, H. H. (1997). How money matters: The effect of school district spending on academic achievement. *Sociology of Education*, 70, 221-237.
- White, K. R. (1985-86). Efficacy of early intervention. *The Journal of Special Education*, 19, 400-416.

- Word, E., Johnson, J., Bain, H. P., Fulton, D. B., Boyd-Zaharias, J., Lintz, M. N., Achilles, C. M., Folger, J., & Breda, C. (1990). *Student/teacher achievement ratio (STAR): Tennessee's K-3 class-size study*. Nashville, TN: Tennessee State Department of Education.
- Yen, W. M. (1986). The choice of scale for educational measurement: An IRT perspective. *Journal of Educational Measurement*, 23, 299-325.
- Zigler, E., & Styfco, S. J. (1994). Is the Perry Preschool better than Head Start? Yes and no. *Early Childhood Research Quarterly*, 9, 269-287.

JEREMY D. FINN is Professor of Education at State University of New York at Buffalo. His research interests include students and schools at risk, student resilience, educational equity, classroom organization, and statistical methods. He has been conducting research on class size and teacher aides since 1985, when he began as external evaluator for Tennessee's Project STAR.

SUSAN B. GERBER is Associate Director of the Center for the Study of Technology in Education and Adjunct Assistant Professor in the Graduate School of Education at State University of New York at Buffalo. Her research interests include policy- and instruction-related approaches to providing all students with essential learning opportunities.

CHARLES M. ACHILLES is Professor of Education Leadership at Eastern Michigan University in Ypsilanti and author of *Let's Put Kids First, Finally: Getting Class Size Right* (Corwin Press, 1999). He was a principal investigator of Project STAR and has conducted class size research continuously since 1984.

JANE BOYD-ZAHARIAS is the Director of Health & Education Research Operative Services, Incorporated (HEROS, Inc.), a not-for-profit research agency in Tennessee. Prior to joining HEROS, she served as Director of Class Size Studies at Tennessee State University. Boyd-Zaharias is a co-principal investigator of STAR follow-up studies and has authored more than 30 publications and presentations related to class size research.

Research
Databases

Sign In to My EBSCOhost

New Search | View Folder | Preferences | Help

Return to
Consortium Library

◀ 10 of 43 ▶ Result List | Refine Search | Print | E-mail

Save | Add to folder

Folder is empty.

Formats: Citation | HTML Full Text

Choose Language | Translate

Title: *THE LONG-TERM EFFECTS OF SMALL CLASSES IN EARLY GRADES: LASTING BENEFITS IN MATHEMATICS ACHIEVEMENT AT GRADE 9*, By: Nye, Barbara, Hedges, Larry V., Konstantopoulos, Spyros, Journal of Experimental Education, 00220973, Spring2001, Vol. 69, Issue 3
Database: Academic Search Premier

Section: Motivation and Social Processes

THE LONG-TERM EFFECTS OF SMALL CLASSES IN EARLY GRADES: LASTING BENEFITS IN MATHEMATICS ACHIEVEMENT AT GRADE 9

Contents

[Method](#)[Results](#)[Discussion](#)

TABLE 1 Third-Grade
Achievement Test
Scores for Students
Present (Stayers)
and Not Present
(Leavers) at Grade 9:
Small Class in Grade
3

TABLE 2 ANOVA
Summary on the
Relationship Between
Math Achievement
and Small Class at
Grade 3

TABLE 3 Ninth-
Grade Average Math
Test Scores, by
Gender, Race, SES,
and Grade 3 Class
Size

TABLE 4 Third-Grade
Achievement Test
Scores for Students
Present (Stayers)
and Not Present

ABSTRACT. Reducing class size to increase academic achievement is a policy option currently of great interest. Although the results of small-scale randomized experiments and some interpretations of large-scale econometric studies point to positive short-term effects of small classes, some scholars view the evidence as ambiguous. **Project STAR** in Tennessee—a 4-year, large-scale randomized experiment on the effects of class size—provided persuasive evidence that small classes have immediate positive effects on academic achievement. Unlike most other early education interventions, these effects persisted for several years after the children returned to regular-sized classes. The authors of the present article report analyses of a 6-year follow-up of the students in that experiment. Class-size effects persisted for at least 6 years and remained large enough to be important for educational policy. The results suggest that small classes in early grades have lasting benefits and that those benefits are greater for minority students than for White students.

Key words: class size, longitudinal study, long-term effects, randomized experiments

CLASS-SIZE REDUCTION is a policy option that is gaining increasing attention nationwide. Some 18 states have recently adopted policies that reduce class size with the goal of improving achievement. Class-size reduction was also included in President Clinton's education initiative.

Until recently, there was no consensus about whether class-size reduction led to increases in academic achievement and other desired outcomes of schooling. Even today, there is no universal consensus among scholars on interpreting the evidence on the relationship between class size and academic achievement. However, we argue that important evidence was provided by **Project STAR**, the Tennessee class-size experiment. This experiment provided rather strong evidence that class-size reduction leads to immediate increases in academic achievement in both reading and mathematics, with some evidence of larger

(Leavers) at Grade 9:
Small Class in All
Grades

effects for minority students. It is not clear from the **STAR** findings, however, whether those increases in academic achievement disappear after the students are no longer in small classes or persist.

TABLE 5 ANOVA
Summary on the
Relationship Between
Math Achievement
and Small Class in All
Grades K-3

In the present study, we investigated whether assignment to small classes in the primary grades has lasting effects on mathematics achievement and, if so, whether those effects vary across schools with different characteristics, such as those in the state of Tennessee. Mathematics achievement is often characterized as a gatekeeper for college admission, a critical filter restricting choice of majors, and a significant predictor of overall college success (College Board, 1990).

TABLE 6 Ninth-
Grade Average Math
Test Scores, by
Gender, Race, SES,
and Grades K-3
Class Size

Mathematics achievement has also been found to be an important predictor of other desirable life outcomes, such as earnings (see, e.g., Paglin & Rufolo, 1990).

REFERENCES

Research on the Effects of Class Size

Well over 100 experiments and quasi-experimental studies of the effects of class size have been conducted, each involving assignment of students to smaller or larger classes. This literature was reviewed by Glass and Smith (1979); Glass, Cahen, Smith, and Filby (1982); Hedges and Stock (1983); and Mosteller, Light, and Sachs (1996). Although there is some disagreement about the interpretation of the experimental research (e.g., Educational Research Services, 1980; Slavin, 1984), these syntheses of research on the effects of class size suggest positive effects of class-size reduction on achievement and affect, with the effects becoming larger as classes become smaller.

However, most class-size experiments have been small scale and short term. Consequently, it is difficult to know whether the special circumstances surrounding the experiment caused the effects and whether they would have occurred in a more natural setting. That is, the small-scale experiments may have high internal validity, but it is difficult to know whether they will generalize to other settings.

Econometric work on education production functions (e.g., Hanushek, 1986) provides another means of studying the effects of class size. This research relies on the fact that naturally occurring class sizes vary across schools. It models the relationship between class size and an outcome (usually academic achievement) while controlling for student characteristics such as social class or prior achievement. The considerable number of econometric studies of the effects of class size on achievement have been reviewed elsewhere (see, e.g., Greenwald, Hedges, & Laine, 1996; Hanushek, 1989; Hedges, Laine, & Greenwald, 1994). There is some controversy over the interpretation of the econometric studies. However, although few researchers would question the external validity of econometric studies, randomized experiments are generally thought to have higher internal validity.

An important addition to the research on class size was the Tennessee class-size experiment, **Project STAR** (Student-Teacher Achievement Ratio). This study (called "one of the great experiments in education in U.S. history" by Mosteller et al., 1996, p. 814) has substantially mitigated many of the problems of other class-size research. It is a large-scale experiment in which both teachers and students were randomized into classrooms within each of a broad range of participating schools. Consequently, it has both high internal validity and considerable external validity.

Initially, all Tennessee school districts were asked to participate in **Project STAR**, and about 100 schools had enough students in each grade to meet the size criterion for participation (at least 57 students per grade were necessary to form one small and two regular-sized classes). This size criterion, which was necessary to permit assignment to class types within schools, excluded very small schools from the study. Ultimately, 79 elementary schools in 42 school districts (of the 141 districts in Tennessee) became sites in the **STAR** experiment. Districts agreed to participate for 4 years and to allow site visitations for verification of class sizes, interviewing, and data collection, including extra student testing. They also allowed random assignment of pupils and teachers to class types from kindergarten through Grade 3.

The experiment randomly assigned kindergarten students to small classes (with 13-17 students), larger classes (with 22-26 students), or larger classes with a full-time classroom aide. Teachers were also randomly assigned to classes of different types. These assignments of students and teachers to class

types were maintained through the third grade. Some students entered the study in the first grade or subsequent grades and were randomly assigned to classes at that time.

We based this article on the data collected as part of the Lasting Benefits Study conducted by the Center of Excellence for Research and Policy on Basic Skills at Tennessee State University (Nye et al., 1994). The analyses reported during the original *Project STAR* (Word et al., 1990) and some analyses conducted since that time (e.g., Finn & Achilles, 1990), as well as reports based on those analyses (such as Mosteller, 1995), have important weaknesses. However, both these analyses and more recent (and satisfactory) analyses of data from *Project STAR* (e.g., Krueger, 1999) support the conclusion that small classes in kindergarten through Grade 3 lead to higher academic achievement at Grade 3. The effects of small classes were found to be remarkably consistent across schools, suggesting that small classes benefit students of all types in all kinds of schools (Nye, Hedges, & Konstantopoulos, 2000).

One particularly interesting aspect of earlier analyses of *Project STAR* is the finding that the benefits of small classes were larger for minority students than for White students (e.g., Finn & Achilles, 1990; Krueger, 1999). Although the differential effects were not always statistically significant, class-size effects were larger for minority students than for White students in nearly all of the analyses conducted.

Lasting Benefits of Early Education Interventions

Although a number of early educational interventions have demonstrated immediate benefits, the effects typically disappeared over time so that the achievement of students who received the intervention was no higher than that of those who did not. For example, McKey et al. (1985) found that initial positive effects of Head Start completely faded out 3 years after the intervention. Haskins (1989) reported similar results in a comprehensive review of model preschool and Head Start programs. In an analysis of 300 early intervention programs, White (1985) found a pattern of initial effects that faded out over time, as did Barnett (1992). The Perry Preschool *Project* appears to be a notable exception to these findings, producing lasting gains in achievement (see, e.g., Barnett, 1985). However, research on many other early interventions suggests that achievement effects disappear by Grade 8 (Barnett, 1992), although Barnett (1995) offered a somewhat more optimistic view of possible lasting effects of early interventions.

There is relatively little evidence about whether small classes in early grades have enduring effects on later academic achievement. However, the Lasting Benefits Study, which tracked students from *Project STAR* after the conclusion of the experiment, did find enduring benefits through Grade 7 (Nye et al., 1994). Subsequent analysis of that data also revealed that students from small classes in the early grades had higher achievement in reading and mathematics as late as Grade 8 (Nye, Hedges, & Konstantopoulos, 1999).

The Present Article

In this article, we address the question of the long-term effects of small classes by examining the achievement of *Project STAR* students when they were in Grade 9, 6 years after the experiment had ended. The use of the *Project STAR* data to examine long-term effects of small classes was complicated by at least two factors.

The first and undoubtedly most important factor is attrition. Even though the students and teachers were initially assigned to classes at random, differences in patterns of attrition between small and regular-sized classes could bias the results of long-term follow-up comparisons of students initially in small and regular-sized classes. Of course, attrition is a problem that vexes not only *Project STAR* but nearly every longitudinal field study. (The possible effects of attrition within *Project STAR* before Grade 3 were considered in a separate article and were found to be negligible, Nye et al., 2000.)

The second complicating factor is that, because new students were enrolled (and randomized) each year, all students did not spend the same amount of time in the experiment. That is, some of the students were part of the experiment for only 1 year (e.g., Grade 3), whereas others were part of the experiment for 4 years (K-3). If the effects are at all cumulative, then one would expect the effects on students with 4 years of small classes to be larger than those on students with only 1 year of small classes.

To address these complicating factors, we carried out several analyses to discern the effects of each of them. We first present an analysis in which the treatment is defined as class-size type (small or regular) at Grade 3. This analysis included students who entered the experiment at any point between kindergarten and Grade 3 and who received at least 1 year of the small (or regular) class size. This type of analysis might be considered typical for a longitudinal experiment.

To examine the effects of a greater or lesser number of years of treatment, we also present the results of analyses in which the treatment was defined as being in small classes for all 4 years (K-3) of the experiment. By comparing the results of this analysis with those of the first analysis, we gained insight about whether treatment effects were larger for students who had received more years of small classes. We used this analysis instead of an analysis by number of years in small classes because the sample sizes for each number of years (i.e., 1, 2, 3, or 4) in small classes were too small for meaningful analyses of class-size effects. Because the pattern of attrition was slightly different in the two analyses, we examined the pattern of differential attrition between treatment groups for each analysis separately.

Method

The design of the **STAR** experiment involved randomly assigning students and teachers to treatments within schools. The statistical analysis was a 2 x 2 x 2 x 2 (Treatment: small/regular-sized classes x Gender: male/female x Race: minority/White x Student Socioeconomic Status: free-lunch eligible/not eligible at Grade 3) analysis of variance (ANOVA), with socioeconomic status (SES) as the covariate. Note that, because our analyses revealed no effects of the full-time classroom aides in these data, we have not included terms for the effects of aides in this analysis. Therefore, the regular-sized class group included students in regular-sized classes both with and without a classroom aide. The dependent variable was the comprehensive mathematics achievement tests administered by the state of Tennessee to ninth graders. To express the magnitude of treatment effects only, we standardized the outcome variable so that the overall variance in the entire sample was 1. Thus, the small-class effects estimated are interpretable as the effect of being in a small class, expressed in standard deviation units.

Results

The Effects of Having Small Classes in at Least One Grade From K-3

In the first analysis of the experiment, we defined the treatment group to include all students who participated in the experiment for 1 or more years and were assigned to a small class in Grade 3. The treatment group therefore included students who had received 1 or more years of small classes. The control group consisted of students who had participated in the experiment for 1 or more years and had been assigned to a regular-sized class in Grade 3.

Attrition from the study. As expected in any longitudinal study, some students who began the study did not remain in the schools studied or were absent when the achievement tests were given. Attrition can be a source of bias in estimating treatment effects if students who leave one treatment group before follow-up are systematically different from corresponding students in other treatment groups. By the ninth grade, follow-up data could not be obtained on more than one half of the students who were in the **STAR** study in Grade 3. Table 1 contains the average SAT mathematics and reading achievement test scores for students who had participated in Grade 3 of the **STAR** experiment, broken down by whether the ninth-grade test data were available. Those whose test scores were available (the stayers) provided the data analyzed in this study. The students whose ninth-grade data were not available were the leavers whose characteristics might have biased the results.

For example, 1,181 (60.9%) of the 1,938 third graders assigned to small classes (who had third-grade math test scores) left the study before they were tested in the ninth grade. The average achievement of the students who left small classes was higher than that of the students who left regular-sized classes (see Table 1). The difference between the mean achievement of the leavers and that of the stayers who had been in small classes was about the same as the difference between the mean achievement of the leavers and that of the stayers who had been in the regular-sized classes.

The third-grade achievement gap between small and regular-sized classes--the crude treatment effect at third grade--was larger among the leavers than among the stayers. The differences between these gaps were statistically significant for both reading and mathematics achievement. Thus, it appears that the treatment effect at third grade was larger among the leavers than among the stayers. That achievement at third grade was positively correlated with achievement at ninth grade suggests that differential attrition may have led to underestimation of the effect of small classes on achievement. That is, it is not plausible that attrition between third and ninth grades made small classes look more favorable.

Treatment effects. The results of the ANOVA are summarized in Table 2, and the factor level means are given in Table 3. The effect of small classes was statistically significant, $F(1,3545) = 12.701$, $p < .001$, and positive, with a treatment effect of about .146 standard deviation units. That is about 97% as large as the effect on mathematics achievement obtained at the end of third grade (which was about .150 standard deviations; see Nye et al., 2000).

The Minority x Small Class interaction was also statistically significant, $F(1,3545) = 9.643$, $p < .01$, indicating that the small-class effect was larger for minority students than for White students. The small-class effect was .186 standard deviations for minority students and only .002 standard deviations for White students. None of the other interactions were statistically significant.

Mathematics achievement was significantly higher for girls than for boys, $F(1,3545) = 16.499$, $p < .01$. The difference in mathematics achievement favoring girls is surprising. In some earlier grades, the girls in **Project STAR** had higher mathematics achievement than the boys (although the differences were not significant). Our result may represent a widening of that gap, but the causes and implications are not clear. As expected, the low-SES and minority students had significantly lower achievement than the other students, $F(1,3545) = 32.844$, $p < .01$, and $F(1,3545) = 5.973$, $p < .05$, respectively.

We carried out a series of other analyses using related analytic models. For example, we eliminated the interaction effects that were not statistically significant and examined slightly different codings of the variables. None of these analyses suggested results that were qualitatively different from those reported here. In each case, the small-class effect had about the same positive magnitude and was statistically significant.

The Effects of Having Small Classes in All 4 Years From K-3

To get some idea of whether longer exposure to small classes produced greater effects than fewer years of small classes, we conducted an analysis in which the treatment group was defined as those who received 4 years of small classes. The comparison group consisted of all other children in the experiment, regardless of whether they had been assigned to the regular-sized class group or had entered the experiment later than kindergarten, had been assigned to the small class group at some point, and had received up to 3 years of small classes.

Attrition from the study. Because the definition of treatment groups in these analyses was quite different from that in the previous analyses, the pattern of attrition need not be the same. Table 4 contains data on the stayers and the leavers. The treatment was defined as 4 years of small classes. The results of these analyses were substantially the same as those of the previous analysis of attrition.

In each analysis, the achievement gap between small and regular-sized classes (the crude treatment effect) was considerably larger among the leavers than among the stayers. The fact that the treatment effect was larger among the leavers than among the stayers suggests that attrition may have led to an underestimate of the treatment effect at ninth grade. Hence, it is not plausible that attrition made small classes look more favorable at ninth grade.

Treatment effects. The results of the ANOVA in which the small-class group was taken to be those who had received 4 years of small classes are summarized in Table 5, and the factor level means are given in Table 6. The effect of small classes was statistically significant, $F(1,3545) = 31.047$, $p < .01$, and positive, with a treatment effect of about .340 standard deviation units. That is considerably larger than the effect estimated in the previous analysis and about 97% as large as the treatment effect in mathematics

achievement obtained at the end of third grade for the students who had experienced 4 years of small classes (which was about .352 standard deviations, see Nye et al., 2000).

The Minority x Small Class interaction was also statistically significant, $F(1,3545) = 4.575$, $p < .05$, indicating that the small-class effect was larger for minority students than for White students. The effect of 4 years of small classes was .207 standard deviations larger for minority students than for White students. None of the other interactions were statistically significant.

As in the previous analysis, girls had significantly higher mathematics achievement than boys, $F(1,3545) = 7.940$, $p < .01$. Low-SES students had significantly lower achievement than other students, $F(1,3545) = 15.051$, $p < .01$. However, although minority students had lower achievement than White students, that difference was not statistically significant, $F(1,3545) = 0.752$, $p > .05$.

It is important to recognize that in this analysis, the students were not randomly assigned to receive 4 (as opposed to 1, 2, or 3) years of small classes. Therefore, the students who received 4 years of small classes, for example, were likely to be different (in ways not entirely accounted for by race, gender, and social class) from those who did not. For example, they may have come from families that were less likely to move. Assuming that students who come from more stable homes may have higher levels of achievement, these students might have been better prepared to benefit from small classes, which could exaggerate the effects. On the other hand, the comparison group against which the students who had received 4 years of small classes were contrasted included substantial numbers of students (at least $n = 835$, or more than 25% of the comparison group) who had received 1 or more years of small classes, which would tend to reduce the apparent effects of small classes.

Taken together, these findings suggest that, although the small-class effect may have been somewhat smaller at the Grade 9 follow up than at the Grade 3 posttest, the effect remained substantial compared with the small-class effect at Grade 3 (no less than 90% of this effect). That is, the small-class effect on achievement may have diminished somewhat, but it definitely did not fade to statistical or practical insignificance after 6 years.

Discussion

The **STAR** experiment demonstrated that small classes led to significantly higher achievement for students in reading and mathematics at the conclusion of the experiment. Our analyses of the data collected by the Lasting Benefits Study demonstrate that the positive effects of small classes in early grades resulted in mathematics achievement gains that persisted at least through Grade 9. Moreover, the achievement effects were similar in magnitude to those observed at Grade 3. That is, the effects of small classes in kindergarten through Grade 3 on achievement did not appear to be fading out by Grade 9.

Lasting benefits were found for all kinds of students in all kinds of schools, but the benefits of small classes were significantly larger for minority students. Thus, small classes may be a way to benefit all students while reducing the gap in achievement between White and minority students.

This 6-year follow-up study was also subject to substantial attrition, but the students who dropped out of small classes actually had higher achievement than those who dropped out of regular-sized classes, suggesting that the observed differences in achievement between students who had been in small and regular-sized classes was not due to attrition.

The **STAR** study provided an important (and perhaps the strongest) piece of converging evidence about the effectiveness of small classes in promoting achievement. The present study provides evidence that these effects do not disappear over time but confer on students from small classes in early grades achievement benefits that last at least until high school. This evidence points to the positive effects of small classes on achievement that are large enough and of sufficient duration to support policies of class-size reduction resulting in small classes (15-17 pupils) in the primary grades.

This study also demonstrates that students who experienced more years of small classes in kindergarten through Grade 3 had higher levels of achievement 6 years later than students who had fewer years of

small classes. Although the evidence that the effects of small classes compound is not definitive, it is strongly suggestive. Therefore, it is likely that the lasting benefits increase with the number of years in small classes.

The mechanism by which small classes lead to higher achievement, and differentially higher achievement for minority students, is not clear. There are, however, some obvious hypotheses. First, small classes may permit teachers to more effectively individualize instruction. For example, with small classes teachers could identify and remedy incipient problems among students at risk for low achievement. Second, small classes may make existing instruction more effective. For example, small classes may lead to fewer disruptions or more effective whole-class instruction.

Understanding the mechanism could lead to more effective ways to implement class-size reductions and to improve their effectiveness. Such understanding is obviously desirable. Similarly, the effect of small changes in class size or of altering the teacher-pupil ratio (without necessarily changing class size) is not immediately obvious from this experiment.

This research has not even answered all the important questions about the effects of small classes on achievement and other desirable outcomes of schooling. For example, the effects of small classes on achievement growth are not yet clear. Although it appears that students in small classes experience an increase in achievement that persists over time, the growth trajectories of students from small and regular-sized classes have not been studied in detail. This too would be desirable.

Other questions concern the effects of small classes on the other outcomes of schooling such as student engagement, motivation, and persistence. Of particular concern is whether students who are in small classes in early grades will be more likely to graduate from high school, more likely to go to college, and more likely to successfully participate in the labor force. Researchers may use data that are continuing to be collected by the Lasting Benefits Study to ultimately answer some of these questions.

Address correspondence to Larry V. Hedges, Department of Education, The University of Chicago, 5835 S. Kimbark Avenue, Chicago, IL 60637. E-mail: hedge@cicero.src.uchicago.edu

TABLE 1 Third-Grade Achievement Test Scores for Students Present (Stayers) and Not Present (Leavers) at Grade 9: Small Class in Grade 3

Legend for Chart:

- A - Grade 9 status
- B - Small class n
- C - Small class Average
- D - Regular-sized class n
- E - Regular-sized class Average
- F - Difference

	A	B	C	D	E	F
Math achievement at Grade 3						
Present (stayers)		757	609.54	1,726	607.13	2.41
Not present (leavers)		1,181	631.56	2,418	621.76	9.80
Difference						-7.39 (*)
Reading achievement at Grade 3						
Present (stayers)		755	608.08	1,703	604.24	3.84
Not present (leavers)		1,160	629.53	2,387	618.92	10.61
Difference						-6.77 (*)

(*) $p < .001$.

TABLE 2 ANOVA Summary on the Relationship Between Math Achievement and Small Class at Grade 3

Legend for Chart:

A - Source
B - df
C - SS
D - MS
E - F
F - Pr > F

A	B	C E	D F
Female students	1	16.021 16.499	16.021 $p < .0001$
Minority	1	5.800 5.973	5.800 $p < .05$
Low SES at Grade 3	1	31.893 32.844	31.893 $p < .0001$
Small class at Grade 3	1	12.333 12.701	12.333 $p < .001$
Small Class x Female interaction	1	0.567 0.584	0.567 --
Small Class x Minority interaction	1	9.364 9.643	9.364 $p < .01$
Small Class x Low SES at Grade 3 interaction	1	2.992 3.081	2.992 --
Within groups	3,545	3,442.341	0.971
Total	3,552	3,521.311	

TABLE 3 Ninth-Grade Average Math Test Scores, by Gender, Race, SES, and Grade 3 Class Size

Legend for Chart:

A - Variable
B - Small class in Grade 3 M
C - Small class in Grade 3 n
D - Regular-sized class in Grade 3(a) M
E - Regular-sized class in Grade 3(a) n
F - Overall M
G - Overall n

	A	B	C	D	E	F	G
Gender							
Male		59.59	379	58.60	1,413	58.79	1,792
Female		62.88	383	60.59	1,387	61.09	1,770
Race							
Minority		60.38	342	57.16	1,316	57.82	1,658
White		61.83	419	61.79	1,476	61.80	1,895
SES							
Low		58.84	438	57.10	987	57.64	1,425
High		64.37	324	60.94	1,813	61.46	2,137
Math achievement		61.19	762	59.59	2,800	59.93	3,562

(a) The referent group consisted of students who were in regular-sized classes in Grade 3 and who were part of **Project STAR** in previous grades.

TABLE 4 Third-Grade Achievement Test Scores for Students Present (Stayers) and Not Present (Leavers) at Grade 9: Small Class in All Grades

Legend for Chart:

- A - Grade 9 status
- B - Small class in all grades n
- C - Small class in all grades Average
- D - Control group n
- E - Control group Average
- F - Difference

	A	B	C	D	E	F
Math achievement at Grade 3						
Present (stayers)		305	614.82	2,178	606.89	7.93
Not present (leavers)		509	614.67	3,090	622.23	19.44
Difference						-11.51(*)

Reading achievement at Grade 3

Present (stayers)	304	615.22	2,154	604.03	11.19
Not present (leavers)	502	639.81	3,045	619.52	20.29
Difference					-9.10(*)

(*) $p < .001$.

TABLE 5 ANOVA Summary on the Relationship Between Math Achievement and Small Class in All Grades K-3

Legend for Chart:

A - Source
 B - df
 C - SS
 D - MS
 E - F
 F - Pr > F

A	B	C E	D F
Female students	1	7.685 7.940	7.685 p < .01
Minority	1	0.728 0.752	0.728 --
Low SES at Grade 3	1	14.567 15.051	14.567 p < .0001
Small class in Grades K-3	1	30.049 31.047	30.049 p < .0001
Small Class x Female interaction	1	0.264 0.273	0.264 --
Small Class x Minority interaction	1	4.428 4.575	4.428 p < .05
Small Class x Low SES at Grade 3 interaction	1	1.136 1.174	1.136 --
Within groups	3,545	3,431.082	0.968
Total	3,552	3,489.939	

TABLE 6 Ninth-Grade Average Math Test Scores, by Gender, Race, SES, and Grades K-3 Class Size

Legend for Chart:

A - Variable
 B - Small class in K-3 M
 C - Small class in K-3 n
 D - Regular-sized class in K-3(a) M
 E - Regular-sized class in K-3(a) n
 F - Overall M
 G - Overall n

A	B	C	D	E	F	G
Gender						
Male	63.12	137	58.43	1,655	58.79	1,792
Female	66.76	156	60.54	1,614	61.09	1,770
Race						

Minority	64.60	110	57.34	1,548	57.82	1,658
White	65.33	183	61.42	1,712	61.80	1,895
SES						
Low	62.48	145	57.09	1,280	57.64	1,425
High	67.59	148	61.01	1,989	61.46	2,137
Math achievement	65.06	293	59.47	3,269	59.93	3,562

(a) The referent group consisted of students who had participated in **Project STAR** and who were not in small classes in all Grades K-3.

REFERENCES

- Barnett, W. S. (1985). Benefit-cost analysis of the Perry Preschool Program and its policy implications. *Educational Evaluation and Policy Analysis*, 7, 333-342.
- Barnett, W. S. (1992). Benefits of compensatory preschool education. *Journal of Human Resources*, 27, 279-312.
- Barnett, W. S. (1995). The long term effects of early childhood programs on cognitive and school outcomes. *The Future of Children*, 5, 25-250.
- The College Board. (1990). *Changing the odds*. New York: Author.
- Educational Research Services. (1980). *Class size research: A critique of recent meta-analysis*. Arlington, VA: Author.
- Finn, J. D., & Achilles, C. M. (1990). Answers and questions about class size: A statewide experiment. *American Educational Research Journal*, 27, 557-577.
- Glass, G. V., & Smith, M. E. (1979). Meta-analysis of research on class size and achievement. *Educational Evaluation and Policy Analysis*, 1, 2-16.
- Glass, G. V., Cahen, L. S., Smith, M. L., & Filby, N. N. (1982). *School class size: Research and policy*. Beverly Hills, CA: Sage.
- Greenwald, R., Hedges, L. V., & Laine, R. (1996). The effect of school resources on student achievement. *Review of Educational Research*, 66, 361-396.
- Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *The Journal of Economic Literature*, 24, 1141-1177.
- Hanushek, E. A. (1989). The impact of differential school expenditures on school performance. *Educational Researcher*, 18(4), 45-65.
- Haskins, R. (1989). Beyond metaphor: The efficacy of early childhood education. *American Psychologist*, 44, 274-282. Hedges, L. V., Laine, R. D., & Greenwald, R. (1994). Does money matter? A meta-analysis of the effects of differential school inputs on student outcomes. *Educational Researcher*, 23(3), 5-14.
- Hedges, L. V., & Stock, W. (1983). The effects of class size: An examination of rival hypotheses. *American Educational Research Journal*, 20, 63-85.

Krueger, A. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics*, 114, 497-531.

McKey, R. H., Condelli, L., Ganson, H., Barret, B. J., McConkey, C., & Plantz, M. C. (1985). *The impact of Head Start on children, families, and communities*. Washington, DC: U.S. Government Printing Office. (DHHS Publication No. OHDS 85-31193)

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

Mosteller, F., Light, R. J., & Sachs, J. A. (1996). Sustained inquiry in education: Lessons learned from skill grouping and class size. *Harvard Educational Review*, 66, 797-842.

Nye, B. A., Hedges, L. V., & Konstantopoulos, S. (1999). The long term effects of small classes: A five year followup of the Tennessee class size experiment. *Educational Evaluation and Policy Analysis*, 21, 127-142.

Nye, B. A., Hedges, L. V., & Konstantopoulos, S. (2000). The effects of small classes on academic achievement: The results of the Tennessee class size experiment. *American Educational Research Journal*, 37, 123-151.

Nye, B. A., Zacharias, J. B., Fulton, B. D., Achilles, C. M., Cain, V. A., & Tollett, D. A. (1994). *The lasting benefits study: A continuing analysis of the effect of small class size in kindergarten through third grade on student achievement test scores in subsequent grade levels. Seventh grade (Tech. Rep.)*. Nashville: Tennessee State University, Center of Excellence for Research in Basic Skills.

Paglin, M., & Rufolo, A. M. (1990). Heterogeneous human capital, occupational choice, and male-female earnings differences. *Journal of Labor Economics*, 8, 123-144.

Slavin, R. (1984). Meta-analysis in education: How has it been used? *Educational Researcher*, 13(8), 6-15, 24-25.

White, K. R. (1985). The efficacy of early intervention. *The Journal of Special Education*, 19, 401-416.

Word, E., Johnston, J., Bain, H., Fulton, D. B., Boyd-Zaharias, J., Lintz, M. N., Achilles, C. M., Folger, J., & Breda, C. (1990). *Student/teacher achievement ratio (STAR): Tennessee's K-3 class-size study*. Nashville, TN: Tennessee Department of Education.

~~~~~

By Barbara Nye, Tennessee State University; Larry V. Hedges and Spyros Konstantopoulos, The University of Chicago

Copyright of **Journal of Experimental Education** is the property of Heldref Publications and its content may not be copied or e-mailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or e-mail articles for individual use.

Source: *Journal of Experimental Education*, Spring2001, Vol. 69 Issue 3, p245, 13p

Item: 4476994

[Top of Page](#)

Formats:  Citation  HTML Full Text

◀ 10 of 43 ▶ [Result List](#) | [Refine Search](#)  Print  E-mail

 Save  Add to folder

 Folder is empty.

© 2005 EBSCO Publishing. [Privacy Policy](#) - [Terms of Use](#)

**State Legislatures Magazine: June 1998**

Editor's note: These articles appeared in the June 1998 issue of NCSL's magazine, *State Legislatures*. To order copies of the magazine or to subscribe, contact the marketing department at (303) 364-7700.

Smaller = Better  
Does It Work? The Research Is Mixed

---

**Smaller = Better?**

A wave of popular support has buoyed up the idea of reducing school class sizes across the country. But does it work? Maybe so, maybe not.

---

By Julie Davis Bell

Reducing the size of elementary school classes is one of the hottest education reforms to hit state legislative agendas. Nationwide, schools average 23 students per class—too large, some say, to provide the kind of individualized attention students really need to succeed. Using research as a guide, supporters are suggesting that states should focus on creating smaller classes. Now states are figuring out what's involved in getting from here to there. It's a story of huge political potential and huge price tags.

It's hard to find many public policy proposals more popular than reducing class size. A March 1997 *Wall Street Journal* poll found that 70 percent of adults believe that reducing class size would result in big improvements for public schools. A 1997 *Education Week* survey found that 83 percent of teachers and 60 percent of principals believed classes should not exceed 17 students. Parents say their children are happier and learn more in smaller classes. Teachers report they have fewer discipline problems, are able to give students more individual help and can cover material faster.

Such strong public opinion numbers rarely surface. Naturally, policymakers everywhere are paying attention. It seems simply to make intuitive sense—children can get a better education in a smaller class, where the teacher can get to know them better and give them more personalized attention. But as legislators get into the issue, they are finding they must proceed with caution. Reducing the size of classes may be the most expensive education initiative they have ever considered. And the research is not at all clear on whether smaller classes translate into improved learning. So, how far should legislators go with something that is politically popular, very expensive and has uncertain outcomes?

"Lower class size is the argument for the reformation of the public school system," says Michigan Senator Joe Conroy, a passionate supporter of smaller classes. Since 1994, Michigan has funded a pilot program in Conroy's home of Flint. It has cost the state approximately \$6 million. Conroy says the results are significant: "Forty-three percent more fourth graders are passing the state reading test, and 18 percent more are passing the state math test." Conroy visits one school building a week in his district. He says he has observed a "sea change" in the way classes are conducted, including greater decorum and teacher enthusiasm. "The train is on the track for more to come," says Conroy. This year he helped to develop the next step for Michigan—a \$20 million program for statewide implementation of class size reduction in the neediest districts.

Michigan is among 20 states that have passed or are considering passing bills mandating smaller classes. These programs vary in several ways. Most focus on the early grades—primarily kindergarten through third—but some, like Tennessee's, target all grades. States may prescribe classroom averages that schools and districts must maintain or

may impose a maximum number of students per class. Some states, such as Arkansas and Virginia, prescribe an average or maximum student-teacher ratio rather than an average or maximum class size. This allows schools and districts additional flexibility because they are able to use teacher aides and still meet state requirements.

#### 20 STUDENTS MAXIMUM

Typically states are setting averages in grades K-3 at around 20 students. Nevada has the lowest mandated size, requiring no more than 15 students per class. Several states approach class size reduction by offering fiscal incentives. In Michigan, the state provides 75 percent of the funds for poor districts that reduce classes to an average of 17 students with a maximum of 19. Oklahoma began its effort with a grant program, but found implementation cumbersome. The state has now set a five-year schedule to reduce classes statewide to a maximum of 20 students per class in grades K-6.

President Clinton has also jumped on board, adding momentum to the already popular movement. In his 1998 State of the Union address, he announced a national effort to reduce classes to an average of 18 students in grades one, two and three. To help states accomplish this goal, he vowed support for hiring 100,000 new teachers who have passed state competency tests, and a construction tax to help build and modernize schools. Estimates are that this plan would cost about \$12 billion over seven years and would be an additional incentive for states to reduce class sizes.

If Clinton's plan becomes law, it will help states with the significant cost involved in smaller classes. States considering the move must make a serious budgetary commitment. In states facing increasing school enrollment, the costs will be even greater. In California it is costing an extra \$800 or so per pupil per school year. This doesn't even include the cost of new classrooms, estimated to be \$1.1 billion. The Michigan class size reduction program for FY 1999 provides nearly \$20 million for grants to eligible schools. Tennessee invested about \$600 million between 1991 and 1996 to implement its program. In Philadelphia, Superintendent David Hornbeck unveiled plans to reduce class size in kindergarten through third grade from an average of 27 students to 20 students by the year 2002. Estimates are that the program would require a thousand new teachers at \$50 million a year and 35 new schools at \$470 million.

#### NO MAGIC NUMBER

Policymakers who look to research to help answer questions about the effectiveness of smaller classes will find some general recommendations, but little definitive evidence. Most states are heeding two general conclusions: They are most effective in the early grades, and they may be the most valuable for at-risk students.

How small do classes have to be before results are observed? Most research indicates that classes must drop to no more than 15 to 17 students. Otherwise, there is no evidence that smaller class size results in better student achievement. But few states have tried such small classes.

"The research hasn't yet answered the questions legislators ask," says Mike Kirst, Stanford University professor and co-director of the Policy Analysis Center for Education (PACE), which provides much of California's education policy research. "There's no consensus on the magic number. There is certainly no research to justify reducing all grades. If we know anything, it is that positive effects are observed when classes around 23 to 25 students are reduced to 15 to 17. Positive effects are observed in grades K-2—not even third grade—and positive effects are observed in disadvantaged kids."

Most legislators pushing change in class size know these statistics, but argue that any reduction is a positive move. "In a state like California, a reduction from 30 to 20 students is huge. It's the best we can do," says Assemblywoman Kerry Mazzone, who chairs the Education Committee. The parents, teachers, students and policymakers in California, in fact, are extremely happy over that state's efforts. The research may say that California's new classes are not small enough to make a difference in student achievement, but the classes are definitely small enough to make a difference in public opinion.

Nowhere has class size reduction been such a hit—or as expensive—as in California. The California class size reduction program, passed in 1996, cost nearly \$1 billion—the largest infusion of money into the state's K-12 education system since the tax-limiting Proposition 13 passed in 1978. About three-fourths of the money was used to reduce class sizes and the remainder went to extra classroom space. The program is voluntary for school districts, but those participating

must reduce class size to 20 or fewer in grades K-3. The state provides the full additional per pupil cost of \$800.

Not only was the California program the largest and most ambitious ever, it was implemented in a remarkably short period of time. And it has turned out to be a political bonanza. The 1996 legislation was so popular with the public and voters that everyone has taken credit for it—including the governor, the superintendent of public instruction and legislative leaders—Democrats and Republicans. In a state often marked by its political contentiousness this was quite a story.

Class size reduction has sparked a whole new outlook on education and education reform in California. Over the last 10 years the state's education system has been down in the dumps; the property tax limiting Proposition 13, coupled with a poor economy and rapidly rising school enrollments, caused tremendous strain. California consistently was ranked at the bottom of the 50 states for per pupil spending and achievement. Public confidence in the schools was low.

The original measure moved quickly. Senator Leroy Greene wrote the first piece of legislation in 1996. The proposal was a modest one, targeting one grade a year. But he was concerned about the lack of research supporting whether this was a good investment. "Five months later the governor discovered it, and it became his class size reduction program. He wanted to do it all at once," Greene says.

As a result of the governor's strong endorsement and support from top education leaders, the legislation was signed into law in a few months. The response among parents and teachers was immediately positive. Perhaps the most telling statistic about the program's political success is that 120 legislators listed class size reduction as a major accomplishment in their 1996 campaign brochures.

#### HURDLES

But there were significant hurdles and very little planning time. California immediately needed 26,000 new teachers. There were not enough certified and substitute teachers for the new classes. The Legislature passed emergency certification legislation, enabling districts to draft candidates from a "reserve pool" of about 200,000 people who had valid state licenses. Some teachers were reassigned from upper grades with little, if any, training. Between the reduction program, the nation's fastest growing student enrollment and an aging teacher work force, estimates are that California will need to hire as many as 250,000 teachers over the next 10 years—more than the current total on the job.

Also immediately needed were 18,000 new classrooms. In California, every unused building was tapped, classrooms were reconfigured, libraries, auditoriums and gyms became classrooms and port-able classrooms were put into service. In doing this, California may have already tapped out its low-cost options. The main alternatives left include more portables or new construction. Having already spent about \$200 million on new classrooms, the state estimates the program will top \$1.1 billion.

That's a hefty price tag, particularly since evaluation plans were not put in place nor were long-term strategies for teachers and facilities adequately mapped out. This year, California has addressed some of the shortcomings of the original legislation to make future steps easier. For example, PACE is among several organizations that will conduct a comprehensive evaluation of the effects of class size reduction.

Senator Greene thinks California is moving too fast. He is concerned about the program's popularity and the response by lawmakers to expand the program by reducing class size in fourth and fifth grades and targeting specific middle and high school classes, such as eighth grade writing. Greene thinks the state would be better off looking at more proven reforms, such as strengthening reading, writing and arithmetic. "Here we are spending a billion dollars without knowing what it's buying."

Assemblywoman Mazzoni continues to be a strong proponent. "Class size reduction has been worth every penny. The deal has had a huge effect on public opinion and brought people together. It has changed public perception about schools and brought back children from the private schools. It's brought the need for adequate facilities to the forefront and is helping to secure the needed public support for funding."

Some good information may soon be forthcoming in California through the PACE study. The evaluation will examine the effect of class size reduction on student achievement; educational opportunities for minority students, students with limited English proficiency, special education and poor students; the quantity and quality of teaching, including professional development and instruction; educational facilities; and parallel effects of other educational reforms.

The California study is likely to offer important information not only to policymakers in that state, but to others considering this substantial investment. In the meantime, plenty of personal testimony points out the benefits of smaller classes. One district in California—the San Juan Unified School District in suburban Sacramento — reported that suspensions were down 19 percent and referrals of students with severe learning disabilities to special education programs were down 16 percent. Significant improvement was made in mathematics scores—an increase of 7 percent to 8 percent. Reading scores went up slightly—4 percent. The San Francisco Unified School District also has seen some gains in standardized test scores since reducing class size.

States planning to reduce the size of classes can learn a great deal by observing the successes and problems of those who have gone before them. Most states can expect problems in finding available and qualified teachers. At least in the early stages, states must make some important decisions: Is it better to have a small class with a less qualified teacher or a large class with a highly qualified teacher? Those who argue that an underqualified teacher is never a good choice suggest that states should focus on teacher quality, not class size. Some studies have suggested that teacher training and experience are more important indicators of student achievement than is class size. Teacher groups also argue that smaller classes are meaningless if teachers have not been trained to work with fewer students. States must deal with important issues of equity as well: When teachers are in high demand, the wealthier districts are more likely to be able to pay for the most qualified.

Stanford's Mike Kirst's advice is "target your resources to the very early grades, disadvantaged pupils, 20 or fewer students per class, and a phase-in with ample supply of quality teachers and sufficient space. Deviations from that are hard to justify."

Many policymakers, however, may find that something so popular is much too hard to resist.

Julle Davis Bell is an education expert at NCSL.

©1998, National Conference of State Legislatures. All rights reserved.



---

## Does It Work? The Research Is Mixed

---

Although over 1,100 studies examine the relationship between class size and student achievement, no definitive conclusions have been reached. Positive results have been demonstrated in Tennessee and Wisconsin, but other research finds little connection between student-teacher ratios and student performance, especially when measured against other types of educational reforms.

Critics argue that class sizes have fallen for decades, but a corresponding increase in student performance has not occurred. Further, international comparisons demonstrate little support for class size reduction. The average score of eighth graders in Korea and Japan was significantly higher than the United States in both math and science on the Third International Mathematics and Science Study, yet the average class sizes in those countries for 13-year-olds were 49 and 36 respectively vs. 23 in the United States.

The largest and perhaps most convincing study supporting class size reduction comes from Tennessee as a result of Project STAR (Student-Teacher Achievement Ratio) initiated in 1985. Spurred by a study detailing small class size success in Indiana, the legislature authorized a four-year program and study. A second phase of the project, Lasting

Benefits, began in 1989 and continues to track students from the initial study to determine whether small class benefits persist over time.

Project STAR examined 7,000 students and teachers randomly assigned to either a small class with approximately 15 students, a regular class with approximately 25 students or a regular class with approximately 25 students and a teacher aide. The four-year study found a consistent pattern of students in smaller classes achieving substantially higher scores than students in regular classes, with or without an aide. The lasting benefits study demonstrates the positive effects of these small classes in grades K-3. Higher achievement from small-class pupils has continued through ninth grade, although the disparity between students in small and regular classes has diminished slightly over time.

A new study of the Wisconsin's SAGE (Student Achievement Guarantee in Education) initiative has also yielded positive results. The state passed legislation in 1995 to phase in reductions in classes to 15 students in poor schools. During the 1996-97 school year, average test scores in smaller first grade classes increased 12 percent to 14 percent more than those of students in regular classes. The average gains were even greater for African American males, with total scores increasing by over 40 percent compared with similar students in regular classes.

While class size reductions cannot guarantee better student performance, both the Tennessee and Wisconsin experience along with other studies suggest that successful programs share key characteristics:

- Class size reduction should be concentrated in the primary years, particularly kindergarten through third grade. Tennessee students returning to regular classes as early as fourth grade maintained significantly higher achievement levels.
- Classes should be reduced to fewer than 20 students. Programs that reduce groups to below 20 have been found to be more effective than programs that retain more than 20 students but use teacher aides and other techniques to lower student-teacher ratios.
- Urban students, particularly minority pupils, benefit more than their peers from smaller classes. In Tennessee, inner city (predominantly minority) students also had significantly higher self-concept and third grade motivation scores than other inner city students.
- Class size reduction works best when coupled with professional development opportunities for teachers. Educators should be trained in new teaching techniques that take advantage of smaller class sizes.

Even if the research did demonstrate a clear link between class size and student performance, the question remains whether limiting class size is the smartest investment compared to other education reforms.

When asked about President Clinton's initiative, Chester Finn of the Hudson Institute think tank noted, "For \$12 billion you could retrain today's teachers so they knew their subjects. You could give each of the nation's 2.7 million teachers a \$1,000 tuition grant to go learn math or really effective techniques for teaching reading." A recent study using data from 60 previous research projects supports Finn's claim. Measuring the effect of a \$500 investment per student, it was found that spending on teacher education had the greatest impact on student achievement. Lowering the student-teacher ratio was found to have a smaller effect than increasing teacher education, experience and salaries.

The study examined only the relationship with student performance. Lowering class size also has been linked to other benefits such as positive emotional development, better and more frequent communication between home and school, fewer special education referrals and decreased teacher absenteeism. However, when measured against the costs of limited school facilities and the difficulty of finding and hiring qualified teachers, support for reducing class size is far from overwhelming.

—Eric Hirsch, NCSL

©1998, National Conference of State Legislatures. All rights reserved.

 NCSL Home Page

Visitor counts for this page.



---

© 2006 National Conference of State Legislatures, All Rights Reserved

**Denver Office:** Tel: 303-364-7700 | Fax: 303-364-7800 | 7700 East First Place | Denver, CO 80230 | [Map](#)

**Washington Office:** Tel: 202-624-5400 | Fax: 202-737-1069 | 444 North Capitol Street, N.W., Suite 515 | Washington, D.C. 20001



## The Public School Parent's Network

An Information Source and Resource Guide for Parents


[PSPN Home](#)
[Academics](#)
[Administrative Issues](#)
[Cultural Sensitivity](#)
[Editorials](#)
[Grade Level Issues](#)
[Health Issues](#)
[Homework](#)
[Learning Disabilities](#)
[Parental Involvement](#)
[Parental Support](#)
[Reference Guide](#)
[Safety at School](#)
[School Choice](#)
[Socialization](#)
[Student Support](#)
[Transportation Issues](#)
[Contact Us](#)
[Site Search](#)

[Up](#)
[Class Size](#)
[Dress Codes](#)
[Religious Issues](#)
[Teacher Quality](#)
[Ed-Speak](#)
[Discipline](#)
[Legislative Issues](#)
[State Links](#)
[Year Round](#)
[After School](#)

## Does Class Size Matter in Public Schools?

### PSPN

The answer to that question appears to be a resounding yes. According to The US Dept of Education publication, ***Reducing Class Size, What Do We Know?*** - March 1999, "the pattern of findings drawn from the existing research leads to the following Summary Research Conclusions:"

1. A consensus of research indicates that class size reduction in the early grades leads to higher student achievement. Researchers are more cautious about the question of the positive effects of class size reduction in 4th through 12th grades. The significant effects of class size reduction on student achievement appear when class size is reduced to a point somewhere between 15 and 20 students, and continue to increase as class size approaches the situation of a 1-to-1 tutorial.
2. The research data from the relevant studies indicate that if class size

**Tell a friend  
about  
psparents.net**



At

PTA.org

Redu

 US De  
Offi  
Secon  
Class

National

 Does  
Class



is reduced from substantially more than 20 students per class to below 20 students, the related increase in student achievement moves the average student from the 50<sup>th</sup> percentile up to somewhere above the 60<sup>th</sup> percentile. For disadvantaged and minority students the effects are somewhat larger.

3. Students, teachers, and parents all report positive effects from the impact of class size reductions on the quality of classroom activity.

### **Why Do Smaller Classes Make a Difference? (excerpts)**

The higher student achievement brought about by class size reduction may result from some of the ways in which reducing class size naturally alters the classroom environment. On being assigned to smaller classes, teachers report that the classroom atmosphere is better, that students can receive more individualized attention, and that the teachers have more flexibility to use different instructional approaches and assignments. One unanticipated result of the Burke County reduced class size initiative was that the teachers found themselves with more classroom space to work with, because they were using the same classrooms with smaller numbers of students. Class size reduction also changes the educational opportunities beyond the classroom, insofar as teachers have a larger portion of time to devote to working with each of their students' parents.

Class size reduction changes numerous features of the classroom situation. There are fewer students to distract each other. Each student in a reduced size class gets more attention on average from the teacher, and more time to speak while the others listen. Reduced class size also reduces the level of noise in a class. One theory offered to explain the positive effects of class size reduction on student achievement simply argues that in smaller classes each student receives a larger portion of the educational resources represented by the teacher's instructional time, and consequently, learns more.

Other researchers have drawn attention to the *quality* of teaching in smaller classes, rather than the *quantity*. The SAGE evaluation study used teacher interviews, classroom observation, and other data-gathering techniques to study what happens in smaller classes, and these researchers suggest that students are benefiting from more *individualized* attention. The teachers know each of their students better, and can keep track of how each student is doing on the learning task of the moment. This knowledge enables the teacher to intervene more effectively to help the individual student make progress.

Researchers also have suggested that smaller classes are more likely to be "friendlier" places, where students develop better relationships with their classmates and with the teacher, encouraging students to become more engaged in classroom learning activities. The smaller the class, the harder it is to escape the positive influence of the classroom educational experience. The explanation for why reduced class size is especially beneficial in the early grades may derive from the fact that in the early grades children are learning how to be students in classrooms where the number of people is larger than the number of people in their families and students are learning a new routine. This socialization theory is also consistent with the research

finding in both *Project STAR* and *SAGE* that the largest increase in student achievement occurs in the first year of a student's experience in a smaller class.

The focus on class size reduction in the early grades also suggests that smaller classes represent a preventive, rather than a remedial, approach. If smaller classes help students start off on the right foot in learning how to adjust to the classroom situation and get engaged in learning activities, then students avoid the more difficult educational path of falling behind, finding help, and catching up to their schoolmates.

The question of class size is not simply a matter of less is more. The pattern of research evidence only favors class size reduction if it is substantial and brings the class size below a certain threshold. Reducing class size from 30 to 25, for example, may well have no effect whatsoever. The research evidence from *Project STAR* showed that students in smaller classes with fewer than 18 students did better when compared with students in larger classes. Given the variations among individual students and teachers and the way they interact, it is unlikely that there is a single "magic number" below which class size suddenly produces a beneficial effect. But it is fairly clear that class size must get somewhere below 20 in order to make a real difference.

Reducing class size to below 20 students leads to higher student achievement. However, class size reduction represents a considerable commitment of funds, and its implementation can have a sizable impact on the availability of qualified teachers. Strengthening teacher quality also leads to higher student achievement. There is more than one way to implement class size reduction, and more than one way to teach in a smaller class. Depending on how it is done, the benefits of class size reduction will be larger or smaller.

Teachers do not necessarily change their behavior when they move to smaller classes. In one observational research study, even though the teachers assigned to smaller classes thought they were teaching differently, the independent observers saw no discernable difference in teacher behavior. Research studies suggest that teachers do not just automatically change their behavior to optimize the potential benefits of smaller class size, and there is a considerable body of research that shows that creating substantial changes in teachers' classroom behavior is no easy feat. Both teachers who are used to teaching in larger classes and teachers newly hired due to the hiring needs associated with class size reduction initiatives will need to have professional training and support to enable them to utilize more fully the advantages of smaller classes. It is clear that many schools will face this challenge in the coming years. The best<sup>6</sup> ways to meet that challenge remain to be found.

## Conclusion

Reducing class size to below 20 students leads to higher student achievement. However, class size reduction represents a considerable commitment of funds, and its implementation can have a sizable impact on the availability of qualified teachers. Strengthening teacher quality also leads to higher student achievement. There is more than one way to implement class size reduction, and more than one way to teach in a smaller

class. Depending on how it is done, the benefits of class size reduction will be larger or smaller.

*The information above is provided from excerpts taken from the "The US Dept of Education's publication, **Reducing Class Size, What Do We Know?** - March 1999", To read this article in it's entirety, please [click here](#).*

---

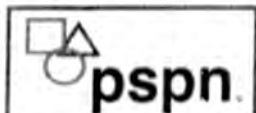
### NEA.org: **Class Size**

(excerpt) While many education reform proposals remain controversial, reducing class size to allow for more individualized attention for students is strongly supported by parents, teachers and education researchers.

NEA supports a class size of 15 students in regular programs and even smaller in programs for students with exceptional needs. Teachers with small classes can spend time and energy helping each child succeed. Smaller classes also enhance safety, discipline and order in the classroom. When qualified teachers teach smaller classes in modern schools, kids learn more. It's common sense, and the research proves it works to increase student achievement. [More>>](#)

---

Copyright (c) 2001-2003 by The Public School Parents Network. All Rights Reserved.  
[Home](#) / [Contact Us](#) / [About Us](#) / [Site Search](#)  
e-mail us at: [info@psparents.net](mailto:info@psparents.net)



The Public School Parents Network



[Key Issues](#)

[Policy Reports](#)

[In the News](#)

[Support Us](#)

[About Us](#)

[Contact Us](#)

[HOME](#)

## Article

[Home](#) > [Key Issues](#) > [Education](#) > [Article](#)

[PRINT ARTICLE](#)

# Smaller isn't better: why reduced class size doesn't improve public education

By: Richard C. Leonardi, posted January 1, 1998

President Clinton called for it in his State of the Union message. The Ohio General Assembly mandated it in last year's school budget. And Ohioans, like most Americans, give it high marks in public opinion polls.

It's class size reduction, and it may be coming to a school near you.

Conventional wisdom suggests that reducing the number of students per teacher would solve a host of education problems, allowing for more personalized attention to students, less classroom disorder, and relief for overworked teachers.

But current research points to a different conclusion. Reducing class size boosts costs dramatically without a matching boost in student achievement. The University of Rochester's Eric A. Hanushek, a leading authority on school finance, recently surveyed national and international school systems and found little or no relationship between class size and achievement test scores. [1]

Hanushek discovered that between 1950 and 1994, student-teacher ratios in U.S. elementary and secondary schools fell by 35%. Nonetheless, math, science and reading scores among American 17 year olds have remained the same throughout most of that time period.[2]

International data showed the same results. When Hanushek examined six international tests in math or science from 70 countries between 1960 and 1990, he learned that, despite very large differences in student-teacher ratios, class size had no significant influence on student performance.[3]

"The econometric evidence is clear. There is little reason to believe that smaller class sizes systematically yield higher student achievement," concluded Hanushek. "While some studies point in that direction, an almost equal number point in the opposite direction." [4]

Ohio's public schools have experienced a similar trend over the past twenty years. There were about four fewer students in the average classroom in 1997 than in 1978, a 22% reduction in student-teacher ratios.[5] Despite these smaller class sizes, Ohio graduation rates have declined and overall proficiency test scores have not improved.

[Advanced Search...](#)

New to the Buckeye  
Sign up for our We  
Digest and membe

## Education Artic

**Viewpoint: Effectiv**

By David W. Kirkp

**Viewpoint: How We  
Pay its Teachers?**

By Matthew Carr

**Viewpoint: Perspec  
Priorities in Educat**

By Matthew Carr

**Testimony: What th  
McKay Program Ca  
About Bringing Sp  
Vouchers to Ohio**

By Greg Forster

**Viewpoint: Dispelli  
Charter Schools in**

By Matthew Carr

**Viewpoint: Ohio's  
Schools are Succe**

By Matthew Carr

**Viewpoint: Charter  
Irresistible Force?**

By David W. Kirkp

[More recent articl](#)

## Education Pres

School Choice Can  
Revitalize Ohio's I  
More...

[More press releas](#)

Economist Richard Vedder and a team of researchers at Ohio University examined thirty different factors for their impact on Ohio public school performance. They found that "the notion that reducing average class size will have meaningful impacts on student learning is, in general, not supported by the evidence." [6]

Despite findings like Hanushek's and Vedder's, policymakers continue to legislate class-size reduction. A recent law passed by the Ohio General Assembly mandated reducing class sizes for some districts down to a 15-to-1 student-teacher ratio. [7] But, because it requires the hiring of more teachers, class size reduction is one of the most expensive reforms a school can adopt—an odd step to take when schools are looking for ways to save money.

One of the problems with reforms like class-size reduction are that they ignore the public school system's underlying, noncompetitive structure. As Hanushek states, "nobody in today's schools has much of an incentive to improve student performance. Careers simply are not made on the basis of student outcomes." [8]

One reform that *does* address and improve upon the existing structure of public schools is gaining ground in Ohio and the nation: child-centered education or "CCE." [9]

CCE proposes that dollars follow students to the schools of their parents' choice. When parents are free to choose where their children are educated, schools will be encouraged to focus on the things that matter most. Nothing matters more to parents than academic performance.

In effect, child-centered education's emphasis on parental accountability and control ensures that education-related careers *would* be made on the basis of student outcomes. Instead of "thinking small" and reducing class sizes, Ohio's policymakers are better advised to think big and expand parental choice in education.

### Notes

[1] Eric A. Hanushek, *The Evidence on Class Size*, Occasional Paper Number 98-1 (Rochester, New York: W. Allen Wallis Institute of Political Economy, University of Rochester, February 1998).

[2] Although Ohio does not participate, the National Assessment of Educational Progress (NAEP) has been given in numerous states since 1970. Hanushek, p. 3.

[3] In fact, Hanushek's research found a positive, but statistically insignificant, relationship between more students per teacher and greater performance. Hanushek, p. 19.

[4] Hanushek, p. 25. Hanushek also challenges the much vaunted Tennessee, or STAR, study. "The inescapable conclusion is that the smaller classes at best matter in kindergarten." He continues: "The STAR data do not suggest that improvements will result from class size reductions at later grades." Hanushek, p. iv.

[5] Teacher-student ratio data were generated from the Ohio Department of Education web site, [http://www.ode.ohio.gov/www/ims/www\\_hist\\_vitals.html](http://www.ode.ohio.gov/www/ims/www_hist_vitals.html).

[6] Richard Vedder, Joshua Hall, and Michael Melander, *Determinants of Ohio Student Performance*, Working Paper (Athens, Ohio: Department of Economics, Ohio University, January 15, 1998), p. 15.

[7] Ohio General Assembly, Ohio House of Representatives, 122<sup>nd</sup> General Assembly, Am. Sub. H.B. 650.

[8] Hanushek, p. v.

[9] For a comprehensive child centered education plan for Ohio's schools, see *Children First: A Discussion Paper on Public School Finance and Education Reform in Ohio* (Dayton, Ohio: The Buckeye Institute

*Richard C. Leonardi, J.D., is a former president of The Buckeye Institute for Public Policy Solutions.*

© 2005 The Buckeye Institute for Public Policy Solutions, All rights reserved.      [Contact Us...](#)      [Privacy a  
Information...](#)

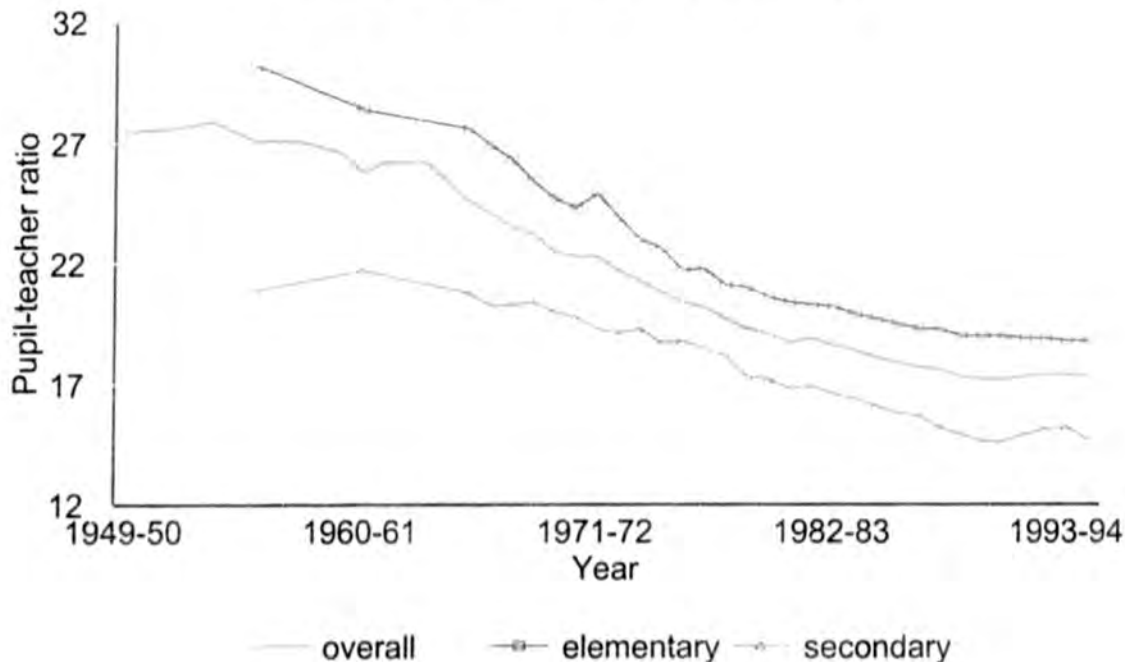
## 1. Basic Aggregate Data

It is common to hear it said that "it is not surprising that achievement is what it is, given the large classes that teachers must face." In reality it is just the history of added teachers without any commensurate increases in student achievement that makes a strong *prima facie* case about the ineffectiveness of class size policies.

There have been consistent and dramatic falls in pupil-teacher ratios over most of the 20<sup>th</sup> century. Figure 1 displays the pattern of pupil-teacher ratios for the period 1950-94. Over this period the overall pupil-teacher ratio fell 35 percent. This decline is the result of steady drops in the pupil-teacher ratio at both the elementary and the secondary school level. The obvious conclusion from this is that, if there is a problem of class size today, there must have been larger problems in the past.

**Figure 1. Pupil-Teacher Ratios**

Overall and elementary and secondary



reductions in pupil-teacher ratios shown in figure 1 do translate into strong effects on spending.

While instructional staff salaries and other spending moved together over the long period, it is also clear from figure 2 that nonsalary spending has grown more rapidly in the past two decades. Thus, the total growth in spending per pupil is not linked in any simple, mechanical way to pupil-teacher ratios, even though increased intensity of instructional staff obviously is an important element.<sup>2</sup>

The other component of the basic aggregate picture is the pattern of student performance. While we do not have representative student achievement data over the entire century, the National Assessment of Educational Progress (NAEP) does provide data since the 1970s.<sup>3</sup> Figures 3-5 display the patterns of NAEP scores in mathematics, science, and reading. These scores are provided for all 17-year-olds and for separate racial and ethnic groups. Three aspects stand out. First, overall performance is approximately the same in 1970 as it is in the mid-1990s. While some differences in patterns exist across the subject areas, the composite picture is one of flat scores over the quarter of a century covered by testing. Second, there has been some convergence of scores between whites and either blacks or Hispanics. For the period up to 1990, the average black-white gap across subject areas narrowed by 0.4 standard deviations (Hauser and Huang 1996), even though the differences remain substantial. Third, the convergence of scores by race and ethnic groups may have stopped during the 1990s. In

---

school systems tended to keep the same numbers of teachers, leading to a decline in the pupil-teacher ratio. With the increase in student population, however, there was no tendency for the pupil-teacher ratio to increase. For just the 1980-90 period, increased intensity of teachers accounted for 34 percent of the growth in total instructional staff spending (Hanushek and Rivkin 1997).

<sup>2</sup>While no systematic analysis is available, it seems plausible that the increased intensity of instructional personnel is directly related to parallel increases in noninstructional personnel. At least a portion of the increase in other costs is undoubtedly attributable to various legal changes including mandates for special education (see below) and desegregation efforts.

<sup>3</sup>A longer time series can be constructed from the Scholastic Aptitude Test (SAT), although using those data introduces some added interpretive issues. The SAT actually fell dramatically from the mid-1960s until the end of the 1970s—suggesting that the achievement decline in the NAEP data neglects an earlier period of achievement fall off. The primary interpretive issue, however, revolves around the voluntary nature of the SAT and the increase in the proportion of high school seniors taking the test. The SAT is taken by a selective group of students who wish to enter competitive colleges and universities. As the proportion taking the test rises, so the hypothesis goes, an increasingly lower achieving group will be drawn into the test, leading to lower scores purely because of changes in test taking. While the exact magnitude of any such effects is uncertain, it seems clear that this change in selectivity has caused some of the SAT decline but not all of it (e.g., see Wirtz et al. 1977; Congressional Budget Office, 1986).

fact, since 1990 there has been a noticeable widening of the racial and ethnic achievement gaps not captured in the calculations above.

The challenge is to reconcile the data on pupil-teacher ratios and resources with the data on student outcomes. On the surface, they suggest that increases in the intensity of teachers and the commensurate increases in spending have had minimal effect on student achievement. But a variety of explanations have been suggested to explain how these data could arise for reasons other than the general ineffectiveness of reduced class size.

**Fig. 3--Mathematics Achievement (NAEP)**  
17-year-olds by race/ethnicity:1973-96

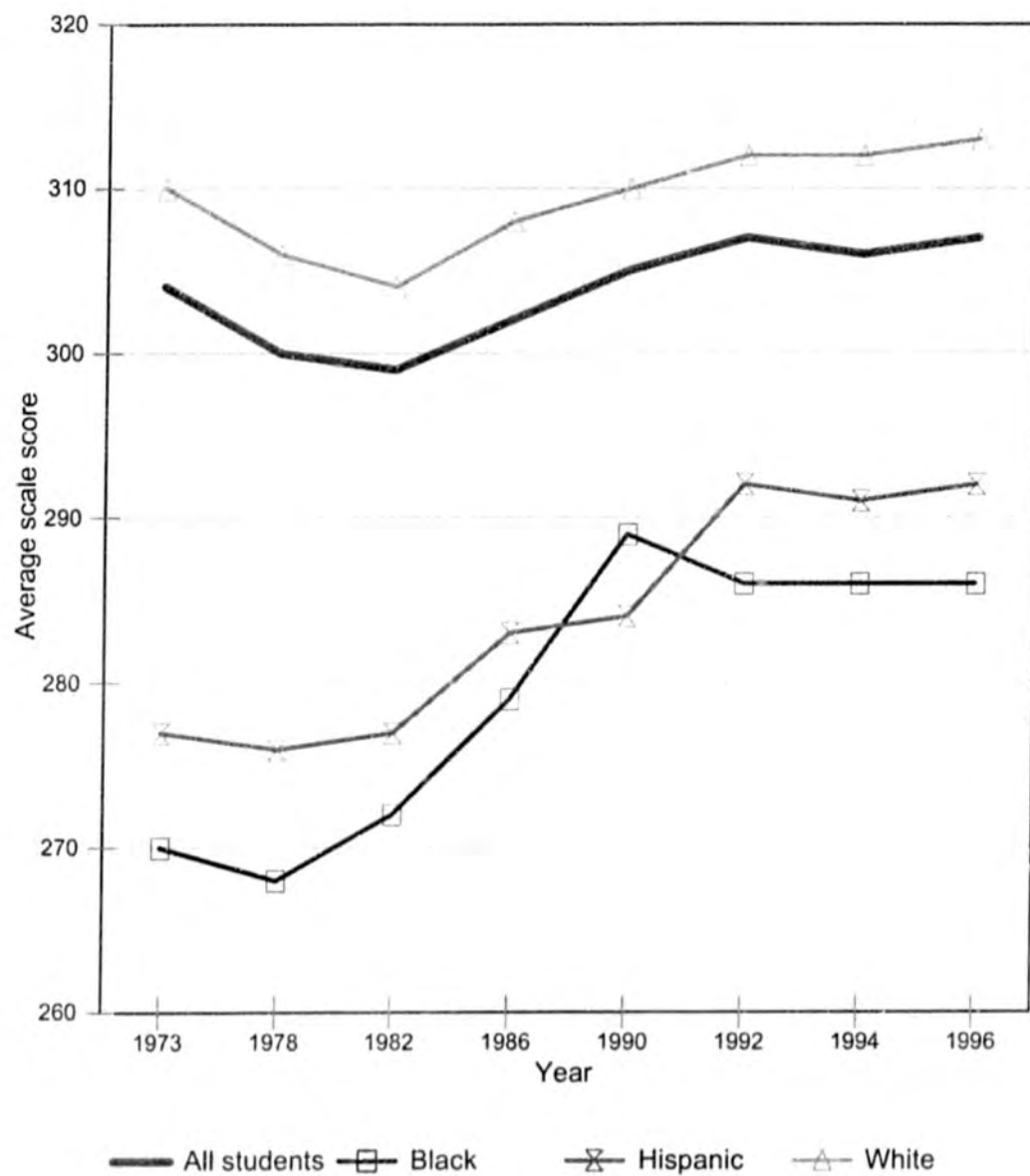


**Fig. 4--Science Achievement (NAEP)**  
17-year-olds by race/ethnicity:1970-96

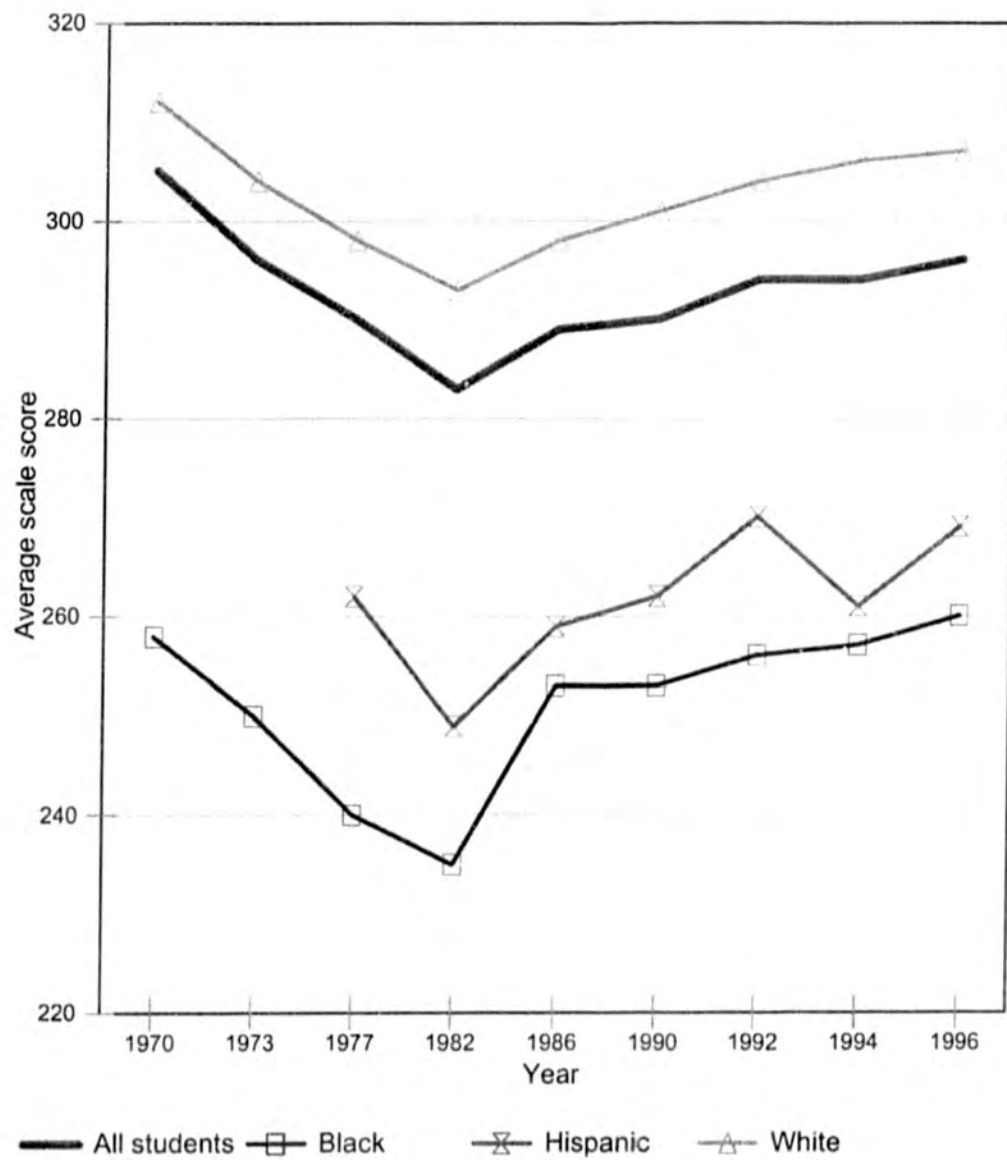


**Fig. 5--Reading Achievement (NAEP)**  
17-year-olds by race/ethnicity: 1971-96

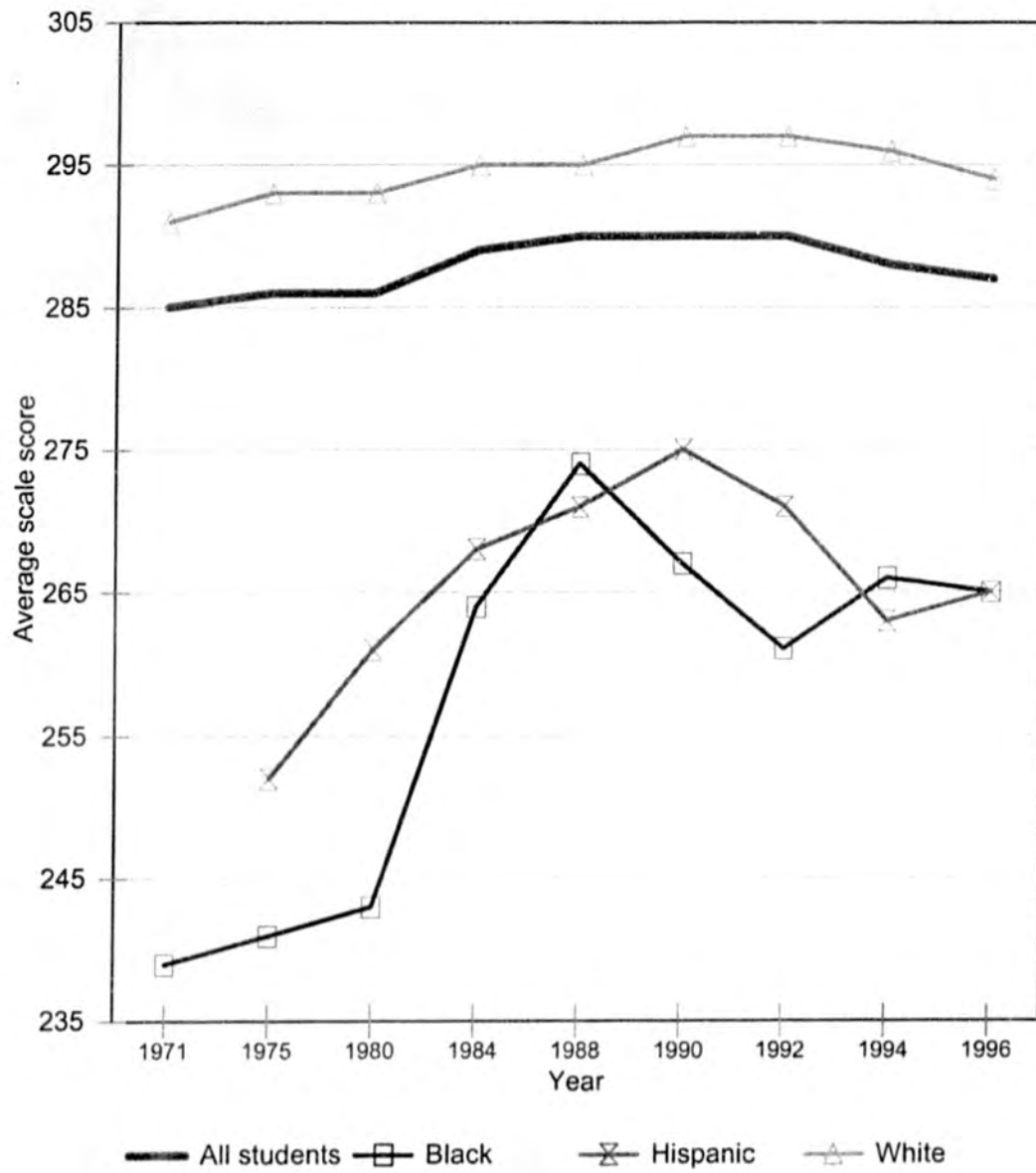


**Table 1. Changes in family characteristics, 1960-1990**

|                                                        | 1970 | 1980 | 1990 |
|--------------------------------------------------------|------|------|------|
| % children in poverty                                  | 14.9 | 17.9 | 19.9 |
| % children under 18 years old living with both parents | 85   | 77   | 73   |
| % high school graduate or more, population age 25-29   | 73.8 | 84.5 | 85.7 |
| % families with children with 3 or more children       | 36.3 | 22.9 | 20.1 |

Source: *U.S. Statistical Abstract, 1992*

#### A. Changes in the student population

One simple explanation for why added resources yield no apparent performance improvement is that students are more poorly prepared or motivated for school, requiring added resources just to stay even. To bolster this view, some point to the increases in children living in single-parent families and the related increases in child poverty rates—both of which are hypothesized to lead to lower student achievement. Table 1 displays these major changes. Between 1970 and 1990, children living in poverty families rose from 14.9 to 19.9 percent, while children living with both parents declined from 85 to 73 percent. But, there have also been other trends that appear to be positive forces on student achievement. As Table 1 also shows, family sizes have fallen, and parental education levels have improved. Over the same period, adults aged 25-29 with a high school or greater level of schooling went from 74 to 86 percent (up from 61 percent in 1960). Moreover, among all families with children, the percentage with three or more children fell from 36 to 20 percent.

It is difficult to know how to net out these opposing trends with any accuracy. Extensive research, beginning with the Coleman Report (Coleman et al. 1966) and continuing through today (Hanushek 1997), has demonstrated that differences in families are very important for student achievement. Most of these studies have not focused their primary attention on families, however, and thus have not delved very far into the measurement and structure of any family influences. Instead, it appears that most have been willing to employ whatever measure of family structure or socio-economic status

might be available. Mayer (1997) suggests that the direct causal impact of family income might be fairly small and that the past works have more identified associations than true causal impacts. This analysis, nonetheless, cannot conclusively indicate whether or not there have been trends in the underlying causal factors (that are correlated in cross-sections with income). Hanushek (1992) indicates that family size may have particularly powerful effects on achievement and indeed may be partly responsible for the narrowing in black-white achievement indicated in Figures 2-4, but again it is difficult to compare the influence of the various trends that have been identified.

Grissmer et al. (1994) attempts to sort out the various factors. That analysis uses econometric techniques to estimate how various family factors influence children's achievement. It then applies these cross-sectionally estimated regression coefficients as weights to the trended family background factors identified above. Their overall findings are that black students performed better over time than would be expected from the trends in black family factors. They attribute this better performance to improvements in schools. On the other hand, white students performed worse over time than would be expected, leading presumably to the opposite conclusion that schools for the majority of students actually got worse over time.

There are again reasons to be skeptical about these results. First, they do not observe or measure differences in schools but instead simply attribute unexplained residual differences in the predicted and observed trends to school factors. In reality any factor that affects achievement, that is unmeasured, and that has changed over their analysis period would be mixed with any school effects. Second, in estimating the cross-sectional models that provide the weights for the trending family factors, no direct measures of school inputs are included. In the standard analysis of misspecified econometric models, this omission will lead to biased estimates of the influence of family factors if school factors are correlated with the included family factors in the cross-sectional data that underlie their estimation. For example, better educated parents might systematically tend to place their children in better schools. In this simple example, a portion of the effects of schools will be incorrectly attributed to the education of parents.<sup>2</sup> Such biased estimates will lead to inappropriate weights for the trended family inputs and will limit the ability to infer anything about the true changes in student inputs over time. Third, one must believe either that the factors identified are the true causal influences (cf. Mayer 1997) or that they maintain a constant relationship with the true causal influences.

---

<sup>2</sup>While it is sometimes possible to ascertain how such statistical misspecification affects the estimated results, the complications here – with multiple factors omitted from the modeling of achievement – make that impossible.

In sum, a variety of changes in family inputs has occurred over time, making it possible that a portion of the increased school resources has gone to offset adverse factors. The evidence is nonetheless quite inconclusive about even the direction of any trend effects, let alone the magnitude. At the same time, the only available quantitative estimates indicate that changing family effects are unable to offset the large observed changes in pupil-teacher ratios and school resources. Indeed, for the nation as a whole, these trends are estimated to have worked in the opposite direction, making the performance of schools appear better than it was. Thus, the most frequently given explanation for the perceived ineffectiveness of historic resource policies does not resolve the puzzle.

#### **B. Special Education and the Changing Structure of Schools**

The discussion until now has focused on pupil-teacher ratios, but pupil-teacher ratios are not the same as class sizes. These data on pupil-teacher ratios reflect the total number of teachers and the total number of students at anytime, not the utilization of these. Take a trivial example. Consider a district that only has two teachers, one of whom spends all day in class with the available students and the other of whom is department head and spends all day evaluating the lesson plans of the classroom teacher. In this case, the pupil-teacher ratio is half that of the class size seen by students. More to the point, if teachers are required to meet fewer classes during the day than the number of classes each student takes, the pupil-teacher ratio will again be less than the average class size. Thus, many people correctly note that typical class sizes observed in schools tend to be larger than the measured pupil-teacher ratio.

The only data that are available over time reflect pupil-teacher ratios. This situation is quite natural, because reporting on actual class sizes requires surveying individual districts about their assignment practices. It is not sufficient to have just the readily available data about numbers of teachers and students. Moreover, class sizes will be influenced by the range of choices given to students and the number of separate courses that individual students are taking. The conceptual ideal behind any measurement is itself not well defined because measuring the actual class sizes faced by students requires a variety of decisions about which classes to count and which not to count. For example, should physical education courses be counted? driver's education? and so forth. This discussion is not meant to imply that we would not like to have data about the varying aspects of schools that influence assignments and class sizes. It merely provides an explanation for why class size data have never been readily available.

Having described the differences, however, it remains to be seen how large the influence of any divergence of class size and pupil-teacher ratios might be on the

aggregate trends previously discussed. In order to influence the trends (as opposed to the observed level during any period), it must be the case that the relationship between pupil-teacher ratios and class sizes is changing over time. While this relationship could change for a variety of reasons from altered work days for teachers to expanded curricular offerings, one possible influence – the increased emphasis on special education – has received the most attention.

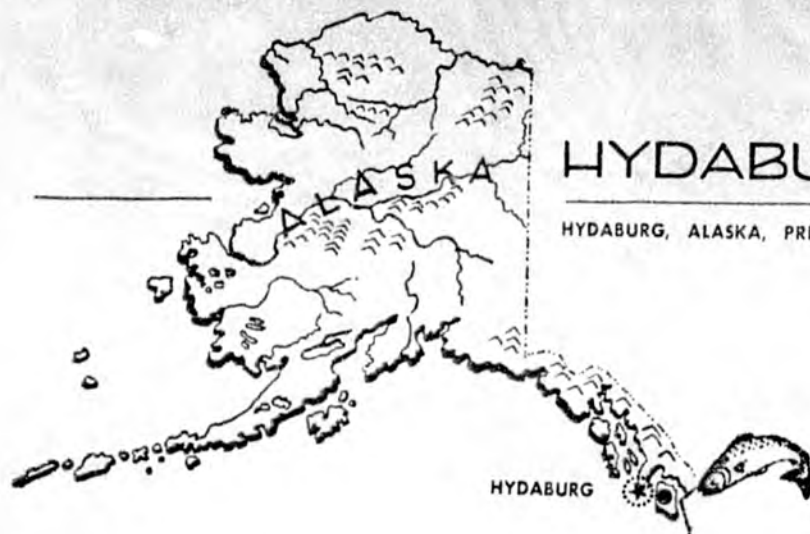
The growth in students with identified handicaps coupled with legal requirements for providing educational services for them has increased the size of the special education sector. Therefore, the expansion of the more staff-intensive special education sector could reduce the overall pupil-teacher ratio without commensurate decreases in regular class sizes. To the extent that mandated programming for handicapped students is driving the fall in the pupil-teacher ratio, regular class sizes may not be declining and, by extension, one might not expect any improvement in measured student performance. This section, which draws on the Hanushek and Rivkin (1997) analysis, provides a simple analysis of the potential importance of special education in explaining the pupil-teacher ratio fall and commensurate increase in educational expenditure.

Concerns about the education of children with both physical and mental disabilities were translated into federal law with the enactment of the Individuals With Disabilities Education Act (IDEA) in 1976.<sup>4</sup> This Act prescribed a series of diagnostics, counseling activities, and services to be provided for handicapped students. To implement this and subsequent laws and regulations, school systems expanded staff and programs, developing entirely new administrative structures in many cases. The general thrust of the educational services has been to provide regular classroom instruction where possible ("mainstreaming") along with specialized instruction to deal with specific needs. The existence of partial categorical funding from outside and of intensive instruction for individual students creates incentives for school systems to expand the population of identified special education students and incentives for parents to seek admission of their children into special education programs (see Hartman[1980], Monk[1990]). The result has been growth of students classified as special education students even as the total student population was falling.

---

<sup>4</sup>While little evidence is available, it is frequently asserted that special education students are not generally included in tests and other measures of performance. Therefore, in assessing performance, it would be appropriate to link expenditure on regular-instruction students with their test performance. On the performance side, however, if a larger proportion of students are identified as special education and if these are generally students who would perform poorly on tests, the shift to increased special education over time should lead to general increases in test scores *ceteris paribus*.

<sup>4</sup>The Act, P.L. 94-142, was originally titled the Education for All Handicapped Children Act.



## HYDABURG CITY SCHOOL

HYDABURG, ALASKA, PRINCE OF WALES ISLAND 99922

February 13, 2006

Representative Eric Croft  
Alaska State Capitol  
Juneau, AK 99801

Dear Rep. Croft

We fully support HB 247 – Time To Teach Act. It would definitely help Hydaburg City School District provide a greater opportunity for our students to get a better foundation for their education.

The district would have the opportunity to keep classes at a level best for learning, with hiring of additional staff.

It is our hope that you will be successful in the passing of HB 247. If we can be of help in this please feel free to give us a call.

Sincerely,

  
Bill Raduenz  
Superintendent



# NEA-ALASKA

*Affiliated with the National Education Association*

**TESTIMONY ON HB 247**  
**HOUSE SPECIAL COMMITTEE on EDUCATION**      **MAY 3, 2005**  
**BILL BJORK, PRESIDENT, NEA-ALASKA**

Good evening, Chair Neuman and members of the House Special Committee on Education. My name is Bill Bjork and I serve as the elected President for the 12,500 public school employees of NEA-Alaska.

I appreciate the opportunity to speak with the committee about HB 247. The #1 factor in student achievement is a qualified teacher in every classroom. The #1 toll for the educator is small class size. Research on the affect of small class size on student achievement has become very clear over the past two decades. The study that led the way to the convincing results is Tennessee's Student/Teacher Achievement Ratio (STAR) project conducted in 1985-1989. NEA-Alaska's letter of support contains a summary of the results.

Let me refer to a study closer to home. In 1993-1996 the "Class Size Reduction Project" was conducted by the University of Alaska Anchorage in four participating elementary schools: one in Juneau and three in Kenai. The project was designed to assess the effectiveness of reduced class size on student achievement and attitude. Significant findings included:

- Student attitude toward learning and school was reported as positive over 90% of the time at all schools during the three years of the project;
- Classroom climate and behavior improved at all participating schools during periods of lower class size;
- Academic achievement improved as a result of reduced class size and instructional innovations; and
- Teachers overwhelmingly reported that small group instruction and challenging learning experiences improved students' self-esteem and attitudes toward school.

Therefore, NEA-Alaska supports HB 247. It allows an individual school to adopt and implement a small class size policy, record the affects on student achievement and report those results to the public. It provides the resources to a school to utilize small class size and, thus, address student achievement in a manner that could reap benefits for several years of each child's education. It also provides an accountability measure that could reap the benefit of public support. Taxpayers will be able to observe the results of the investment in a measurable way.

HB 247 deserves your support for what it does not do. It does not mandate that a school district do this in every school. It allows each school to decide if this is a tool they wish to utilize. It is local control at its best.

I urge your support of HB 247. I appreciate your time and I would be happy to attempt to answer questions.



# NEA-ALASKA

*Affiliated with the National Education Association*

**TESTIMONY ON HB 247  
HOUSE SPECIAL COMMITTEE on EDUCATION  
BILL BJORK, PRESIDENT, NEA-ALASKA**

**MAY 3, 2005**

Good evening, Chair Neuman and members of the House Special Committee on Education. My name is Bill Bjork and I serve as the elected President for the 12,500 public school employees of NEA-Alaska.

I appreciate the opportunity to speak with the committee about HB 247. The #1 factor in student achievement is a qualified teacher in every classroom. The #1 toll for the educator is small class size. Research on the affect of small class size on student achievement has become very clear over the past two decades. The study that led the way to the convincing results is Tennessee's Student/Teacher Achievement Ratio (STAR) project conducted in 1985-1989. NEA-Alaska's letter of support contains a summary of the results.

Let me refer to a study closer to home. In 1993-1996 the "Class Size Reduction Project" was conducted by the University of Alaska Anchorage in four participating elementary schools: one in Juneau and three in Kenai. The project was designed to assess the effectiveness of reduced class size on student achievement and attitude. Significant findings included:

- Student attitude toward learning and school was reported as positive over 90% of the time at all schools during the three years of the project;
- Classroom climate and behavior improved at all participating schools during periods of lower class size;
- Academic achievement improved as a result of reduced class size and instructional innovations; and
- Teachers overwhelmingly reported that small group instruction and challenging learning experiences improved students' self-esteem and attitudes toward school.

Therefore, NEA-Alaska supports HB 247. It allows an individual school to adopt and implement a small class size policy, record the affects on student achievement and report those results to the public. It provides the resources to a school to utilize small class size and, thus, address student achievement in a manner that could reap benefits for several years of each child's education. It also provides an accountability measure that could reap the benefit of public support. Taxpayers will be able to observe the results of the investment in a measurable way.

HB 247 deserves your support for what it does not do. It does not mandate that a school district do this in every school. It allows each school to decide if this is a tool they wish to utilize. It is local control at its best.

I urge your support of HB 247. I appreciate your time and I would be happy to attempt to answer questions.

# FISCAL NOTE

**STATE OF ALASKA**  
**2006 LEGISLATIVE SESSION**

Fiscal Note Number: \_\_\_\_\_

Bill Version: \_\_\_\_\_

HB 247

( ) Publish Date: \_\_\_\_\_

Revision Date/Time (Note if correction): \_\_\_\_\_

Title An Act establishing a grant program to support  
voluntary class size reduction.

Dept. Affected: Education & Early Development

RDU K-12 Support

Component Foundation Program

Sponsor Representative Croft

Requester Health, Education & Social Services

Component No. 141

## Expenditures/Revenues

(Thousands of Dollars)

Note: Amounts do not include inflation unless otherwise noted below.

| OPERATING EXPENDITURES | FY 2007          | FY 2008          | FY 2009  | FY 2010  | FY 2011  | FY 2012  |
|------------------------|------------------|------------------|----------|----------|----------|----------|
| Personal Services      |                  |                  |          |          |          |          |
| Travel                 |                  |                  |          |          |          |          |
| Contractual            |                  |                  |          |          |          |          |
| Supplies               |                  |                  |          |          |          |          |
| Equipment              |                  |                  |          |          |          |          |
| Land & Structures      |                  |                  |          |          |          |          |
| Grants & Claims        | 120,239.1        | 128,980.9        | *        | *        | *        | *        |
| Miscellaneous          |                  |                  |          |          |          |          |
| <b>TOTAL OPERATING</b> | <b>120,239.1</b> | <b>128,980.9</b> | <b>*</b> | <b>*</b> | <b>*</b> | <b>*</b> |

|                             |  |  |  |  |  |  |
|-----------------------------|--|--|--|--|--|--|
| <b>CAPITAL EXPENDITURES</b> |  |  |  |  |  |  |
|-----------------------------|--|--|--|--|--|--|

|                               |  |  |  |  |  |  |
|-------------------------------|--|--|--|--|--|--|
| <b>CHANGE IN REVENUES ( )</b> |  |  |  |  |  |  |
|-------------------------------|--|--|--|--|--|--|

## FUND SOURCE

(Thousands of Dollars)

|                                         |                  |                  |          |          |          |          |
|-----------------------------------------|------------------|------------------|----------|----------|----------|----------|
| 1002 Federal Receipts                   |                  |                  |          |          |          |          |
| 1003 GF Match                           |                  |                  |          |          |          |          |
| 1004 GF                                 | 120,239.1        | 128,980.9        | *        | *        | *        | *        |
| 1005 GF/Program Receipts                |                  |                  |          |          |          |          |
| 1037 GF/Mental Health                   |                  |                  |          |          |          |          |
| Other (Specify Type--Do not abbreviate) |                  |                  |          |          |          |          |
| <b>TOTAL</b>                            | <b>120,239.1</b> | <b>128,980.9</b> | <b>*</b> | <b>*</b> | <b>*</b> | <b>*</b> |

Estimate of any current year (FY2006) cost: 0.0

Mark this box (X) if funding for this bill is included in the Governor's FY 2007 budget proposal: ☐

## POSITIONS

|           |  |  |  |  |  |  |
|-----------|--|--|--|--|--|--|
| Full-time |  |  |  |  |  |  |
| Part-time |  |  |  |  |  |  |
| Temporary |  |  |  |  |  |  |

## ANALYSIS: (Attach a separate page if necessary)

This legislation would establish a grant to provide additional funding above the base student allocation of \$4,919 up to \$8,000. In order to qualify for the grant a school must keep their class size for Kindergarten through 3rd grade at no more than 15 students per classroom.

Requirements are as follows: a 15:1 PTR is maintained over a 5 year period; a study is completed on the academic progress; and a comparative analysis of the progress and outcome of these student's achievement. This fiscal note assumes all school will qualify for the grant.

The department will annually increase the \$8,000 based on the Anchorage Consumer Price Index as determined by the Department of Labor.

[The FY08 projection is based on an average increase of the CPI-U over the last 3 years.]

Prepared by: Eddy Jeans, Director

Division School Finance

Phone 465-8679

Date/Time 3/13/06 1:49 PM

Approved by: Roger Sampson, Commissioner

Agency Education & Early Development

Date 03/13/2006

Rep Neuman

This article address' legislation filed by Eric Croft and is now in your Committee. It points out potential problems that would arise should the legislature decide to grant any monies to districts that shrink pupil-teacher ratio in a classroom.



[Click here to return to the original story](#)

## Croft's bill offers incentives to schools with small classes

Legislation would cost \$150 million a year if all districts cut class sizes

Rep. Eric Croft wants the state to give more money to schools that reduce class sizes to 15 or fewer students.

House Bill 247, introduced this week, would apply to grades kindergarten through three.

Schools would have to agree to keep class sizes that small for five years so the effect on student achievement could be studied.

The bill proposes funding students in those small classes at \$8,000 each. The current base allocation is \$4,576.

"I want teachers to have time to teach," Croft, D-Anchorage, said. "When you have 25 first-graders, you're a traffic cop. You're wiping noses and putting on 25 pairs of boots. When you have 15 students in your class, you can actually teach."

Dave Newton, principal at Auke Bay Elementary in Juneau, said the bill was a good idea as long as the additional funding was guaranteed over time.

At first glance, the new funding would be more than enough to hire teachers and add more specialists to cover the new teachers' preparation time.

But it's not clear if schools would have enough classrooms, and whether the new funding would cover all of the operating costs of new schools, if they were needed.

For example, the bill would allow three classrooms of 25 students to become five classrooms of 15 students.

Over an entire district, such as Juneau, that sort of transformation would mean hiring many new teachers and finding a lot of space.

At Auke Bay Elementary, 10 current classrooms for kindergarten through

grade three, which have 22 to 30 kids in them, could become 17 classrooms of 15 or fewer children.

"The district probably would have to maximize the existing spaces in buildings, which may include bringing in portables until additional schools would be built," Newton said.

"If you want smaller classes, you have to find the space and you have to find the teachers," said Joan McRobbie, a senior policy associate at WestEd, a regional nonprofit that studies education.

Some states that mandated smaller classes blunted the benefits by hiring inexperienced teachers, she said.

In some cases, states didn't give enough money to cover all of the new costs, including adding to facilities, McRobbie said.

In California, which mandated smaller classes in early grades but didn't fully fund them, schools increased classes in the intermediate grades, Newton said.

"We wanted to encourage rather than mandate, and we wanted to provide maximum flexibility," Croft said.

Croft expects that if the bill becomes law, school districts would phase in smaller classes over time. Districts would have enough time to plan new facilities and seek state help with construction costs, just as they do now, he said.

The state hasn't studied the fiscal effects of Croft's bill, which hasn't had a hearing yet. But Croft said it could cost \$150 million a year if every school district implemented it.

"That's big but that's about what we have increased the education money by over the last two years," he said. "It's not off-the-charts fantasy."

Still, the House is stuck this session at raising school funding by about \$70 million, and the Senate agreed to only \$32 million, pending resolution of retirement fund issues.

Rep. Jim Holm, R-Fairbanks, introduced a bill last month that would mandate class sizes of 20 or fewer in kindergarten through grade three. House Bill 228 also hasn't had a hearing.

Click here to return to story:

[http://juneauempire.com/stories/040805/sta\\_20050408005.shtml](http://juneauempire.com/stories/040805/sta_20050408005.shtml)