

The Case for Smaller Classes

and for evaluating what works in the schoolroom

by FREDERICK MOSTELLER



THE UNITED STATES IS NOW ENGAGED IN A LARGE AND EXTENSIVE program to improve our nation's public-school systems. Last year Congress adopted President Clinton's initiative to begin the federal funding necessary to add 100,000 elementary-school teachers and a substantial number of new classroom buildings throughout the country.

Some states are undertaking similar initiatives. An example is the California program, begun in September 1996 under Republican governor Pete Wilson, that aimed to reduce class size in kindergarten through third grade to 20 students. Although California's new reduced-class-size programs have been criticized for inadequate preparation, the state's failure to line up enough additional classrooms and teachers may have encouraged President Clinton's proposals. Meanwhile, three other states—Florida, Georgia, and Utah—have been considering smaller classes in the early grades.

Like many ideas about how to improve or reform education, the effort to reduce class size is controversial. Some critics of public education see the program as a boondoggle, at worst a payoff to entrenched teachers' unions. Others say strengthening teacher training or generally improving the quality of teachers would be more efficacious.

This disagreement over methods of educational reform points to a troubling discovery—we have little, if any, objective, useful data on what really works in education, a quarter-trillion-dollar-plus enterprise that vitally affects our children. In the case of class-size reduction, however, there *are* such data, and they offer strong evidence that smaller classes in the early grades improve children's learning. We need, therefore, to pay attention to this example of an educational experiment, both for its immediate

bearing on the current issue of class-size reduction and for its larger message about how we ought to go about evaluating what works to improve education.

Project STAR (Student/Teacher Achievement Ratio), the state of Tennessee's four-year study of the educational effects of class size and teachers' aides in the early grades, is one of the great experiments in education in U.S. history. Its importance derives in part from its being a statewide study and in part from its size and duration. But even more important is the care taken in the study's design and execution. Not only are the findings valuable, but Project STAR is also extremely important as an example of the kind of experiment needed to appraise other school programs, and as proof that such a project can be implemented successfully on a statewide basis.

IN THE LATE 1980S, THEN-TENNESSEE GOVERNOR LAMAR ALEXANDER (currently a candidate for the Republican presidential nomination) had made education a top priority for his second term. The state legislature and the educational community had been intrigued by a modest-sized Indiana study called Project Prime Time, which found benefits in having small classes in the early grades. The legislature was also aware of an investigation by Gene V. Glass and his colleagues at the University of Colorado and Murdoch University in Australia that used meta-analysis (a way of pooling information from several separate studies to strengthen evidence) to review the literature on the effects of class size. The results of this investigation suggested that a class size of 15 or fewer would be needed to make a noticeable improvement in classroom performance. Meta-analysis, however, was not viewed favorably by all professionals at that time, and

the effect of class size continued to be seriously debated.

Noting the expense associated with additional classrooms and teachers, the Tennessee legislature decided that it would be wise to have a solid research base before adopting such a major program. In addition to studying class size, the legislature wanted to evaluate the effectiveness of adding a teacher's aide to a regular-size class. It therefore authorized and funded Project STAR.

The idea that drove the Tennessee study is that teachers in smaller classes have more time to give to individual children. In addition, teachers and administrators who advocate small classes for students who are beginning school often say they are dealing with a "start-up phenomenon." When children first come to school, they face a great deal of confusion. They need to learn to cooperate with others, to learn how to learn, and to get organized to become students. They arrive from a variety of homes and backgrounds, and many need training in paying attention, carrying out tasks, and engaging in appropriate behavior toward others in a working situation.

The study was carried out in three kinds of groups: small class size (13 to 17 students); regular class size (22 to 25) with a teacher's aide; and regular class size without a teacher's aide. The study began in kindergarten and continued through the third grade. The children moved into regular-size classes in the fourth grade. By comparing average pupil performance in the different kinds of classes, researchers were able to assess the relative benefits of small classes and the presence of a teacher's aide. The experiment involved 79 schools from inner-city, urban, suburban, and rural areas, so that the progress of children from different backgrounds could be evaluated. In all, the experiment involved about 6,400 students during its four years.

As Project STAR approached its final year, the staff requested and received funding for an additional program. The Lasting Benefits Study was designed to follow all three groups of students as they moved into regular-size classes after third grade.

Two kinds of tests were used to assess student performance: standardized tests and curriculum-based tests. Standardized tests have the advantage of being used nationwide, but the disadvantage of not being geared directly to the course of study taught locally. Curriculum-based tests reverse those benefits and disadvantages: they measure more directly the increased knowledge of what was actually taught, but usually cannot tell how the results compare with the national picture.

After four years, it was clear that smaller classes did bring substantial improvement in early learning in cognitive subjects such as reading and arithmetic (for details on methods and findings, refer to the works by Mosteller et al. cited in the box at right). Following the groups further, the Lasting Benefits Study demonstrated that the positive effects persisted into grades 4, 5, 6, and 7, so that students who had originally been enrolled in smaller classes continued to perform better than their grademates who had

started in larger classes. In the first two years of Project STAR, the gains of the minority students (primarily African Americans) were twice as great as those of the majority students; in subsequent years, however, they settled back to about the same gain as the rest. The presence of teachers' aides during Project STAR, though beneficial, did not produce improvements comparable to the effect of the reduction in class size, nor did their presence seem to have as much lasting benefit after third grade.

IS REDUCING CLASS SIZE THE BEST, MOST COST-EFFECTIVE, REFORM? THE Tennessee study does not prove that. Some experts, such as Robert E. Slavin, codirector of Johns Hopkins University's Center for Research on the Education of Students Placed at Risk, focus on style of teaching and teacher quality as more important. But the valid data needed to assess and compare many alternative strategies simply don't exist. For example, we do not have strong evidence about the effectiveness (if any) of the widely used, but much debated, procedure of tracking (breaking classes into groups of comparable attainments).

What we do know from the Tennessee study is that this kind of investment *does* have a beneficial result. After reviewing the Project STAR findings, Tennessee policymakers asked themselves where it would be most effective to introduce this intervention. They decided to implement the small-class program in the 17 school districts where the children seemed most at risk of falling behind—those districts with the lowest per-capita incomes. This change meant decreasing class size in only 4 percent of the classrooms in the state. The results of the first three years of this program, called Project Challenge, have been encouraging. Thanks to the smaller classes, the children from these districts are performing better on both standardized and curriculum-oriented tests than pupils from the same districts in earlier years.

Indeed, their end-of-year performance has raised their district ranking in arithmetic and reading from far *below* the state average for all districts to *above* average.

What we also know from the Tennessee study is that we need more experiments of comparable quality to guide intelligent, effective policymaking for such a huge and vitally important enterprise as education. It seems strange that, after almost a century of educational research, we should be arguing about the outcome of one substantial controlled experiment concerning one classroom feature. I envision collections of districts or states joining together to design studies of mutual interest, just as medical institutions now routinely join together to carry out cooperative randomized clinical trials. The medical and health-care communities have come to expect this. The education community should expect no less. ▢

Frederick Mosteller, LL.D. '91, is professor of mathematical statistics emeritus. This article is in part adapted from, and updates, his analysis of the STAR program in the Harvard Education Letter of July/August 1997.

Extra Credit: Additional Readings

- Frederick Mosteller, "The Tennessee Study of Class Size in the Early School Grades," *The Future of Children*, vol. 5, no. 2 (Summer/Fall 1995): 113-127.
- Frederick Mosteller, "Smaller Classes Do Make a Difference in the Early Grades," *Harvard Education Letter*, vol. 13, no. 4 (July/August 1997): 5-7.
- Frederick Mosteller, Richard J. Light, and Jason A. Sachs, "Sustained Inquiry in Education: Lessons from Skill Grouping and Class Size," *Harvard Educational Review*, vol. 66, no. 4 (Winter 1996): 797-842.
- Robert E. Slavin, *Cooperative Learning*, second edition (Allyn & Bacon, 1995).

