Small classes in grades K-3 boost student academic performance in all subjects and in prosocial behavior. Results are both short- and long-term. One study explored the theory that a major cause behind improved academic achievement involves improved student behavior, which increases student engagement in the classroom. Two other studies provide evidence that teachers spend more time on task and with individual students because of less time spent on dealing with disciplinary problems and interactions not related to learning objectives, compared with teachers in larger classes. This study suggests that, given a reasonable pupil-teacher ratio (PTR), the closer the K-3 actual class sizes are to the building PTR, the better the student outcomes (academic, behavior, and so forth). Small classes can be obtained with little or no added funding if attention is paid to personnel assignments to get PTR and class sizes nearly equal. High-performing and small-class schools are characterized by a minimum of "pullouts" and disruptions; high parental support; and planned and coherent education programs. Appendix A describes the Tennessee STAR project, Appendix B lists some major differences between PTR and class size, and an attachment provides recommendations for class-size change. (Contains 78 references.)
School Improvement Should Rely on Reliable, Scientific Evidence.
Why Did "No Child Left Behind" Leave Class Size Behind?

C. M. Achilles
Jean D. Krieger
J. D. Finn
Mark Sharp

February 2003
SCHOOL IMPROVEMENT SHOULD RELY ON RELIABLE, SCIENTIFIC EVIDENCE. WHY DID “NO CHILD LEFT BEHIND” LEAVE CLASS SIZE BEHIND?

A Symposium Presented to the Conference Within A Convention, American Association of School Administrators (AASA).

February 20-23, 2003
New Orleans, LA

Participants and Authors

C. M. Achilles, EdD
Professor, Education Leadership
Eastern Michigan University (50%)
Seton Hall University (50%)
Mail Address: 4477 Snug Harbor,
Geneva, NY 14456
315-789-2399
e-mail: plato9936@yahoo.com

J. D. Finn, PhD
Professor, Graduate School of Education
Christopher Baldy Hall
SUNY, Buffalo, NY 14260

Jean D. Krieger, PhD
Principal, Woodlake Elem.
Mandeville, LA and Adjunct Professor, U. of N. Orleans

Mark Sharp
Adjunct Professor, Ed. Admin.
304 Porter
E. Michigan University
Ypsilanti, MI 48197

SCIENTIFICALLY-BASED RESEARCH (SBR) SUPPORTED BY EXPLICIT THEORY AND DEMONSTRATED REASONABLE COSTS SHOW HOW TO LEAVE NO CHILD BEHIND. WHY ARE EDUCATORS HESITATING TO IMPLEMENT CLASS-SIZE FINDINGS?

A Conference Paper by C. M. Achilles and J. D. Finn
SCHOOL IMPROVEMENT SHOULD RELY ON RELIABLE, SCIENTIFIC EVIDENCE.
WHY DID “NO CHILD LEFT BEHIND” LEAVE CLASS SIZE BEHIND?

SYMPOSIUM ABSTRACT

Small classes (k-3) boost student academic performance in all subjects and in behavior such as reduced indiscipline (e.g. STAR, SAGE, etc.). Results are both short and long-term. This symposium reviews a) the facts and theories about “why” small classes have positive outcomes, and b) some cost issues in obtaining small classes, K-3.

One study explored the theory that a major “why” of the class-size effect is student behavior, specifically student engagement, that includes both learning and pro/anti-social behavior. Theory and empirical research strongly support the student-behavior hypothesis.

Two studies provide evidence for other sources of the class-size effect. One study included structured observations of verbal and non-verbal interactions between early primary students and teachers judged effective (by independent observations and teaching-record assessments). Teachers (n=11) had small (18 or fewer students) or regular (25 or more students) classes. Teachers in small classes spent more time on task and with individual students. They spent less time on negative events and on interactions not related to learning objectives than did teachers in larger classes.

The third study included seven separate approaches to study differences between class size and pupil-teacher ratio, (PTR): their costs, impacts on students, and policy confusions. The average difference in the U.S. between PTR and class-size is about n= 10 pupils, the difference between the regular classes and the small classes in studies such as STAR, SAGE, and others. This finding explains the confusion generated when policy persons who use PTR data declare that class-size doesn’t matter! Done according to the research, small classes (n=15 or so) and high student performance can be achieved in k-3 at essentially no added costs by using reasonable “trade-offs.” Low-performing schools and districts have greater disparity between actual class sizes and their PTR’s than do higher performing districts. This study and the next one suggest the following hypothesis: Given a reasonable building-level PTR, the closer the k-3 actual class sizes are to the building PTR, the better are student outcomes (academic, behavior, etc.).

Actual processes used in selected districts (approx. 10) to obtain small (n=15) classes, K-3 show that small classes can be obtained with little or no added funding if attention is paid to personnel assignments to get PTR and class sizes nearly equal. High performing and small-class schools are characterized by a minimum of “pullouts” and disruptions; high parental support; and planned and coherent education programs. Given the high achievement in academics and behavior and low costs that can accompany small classes, one must wonder why, with it’s avowed encouragement for solid, replicable research, the “No Child Left Behind” (NCLB) Act left the most solidly researched, replicable education “treatment”—class-size—behind.
SCHOOL IMPROVEMENT SHOULD RELY ON RELIABLE, SCIENTIFIC EVIDENCE. WHY DID "NO CHILD LEFT BEHIND" LEAVE CLASS SIZE BEHIND?

SCIENTIFICALLY-BASED RESEARCH (SBR). DRAFT

C. M. Achilles, EdD and Jeremy D. Finn PhD

Introduction

In this paper we consider the requirement for scientifically based research (SBR) as expressed in the No Child Left Behind (NCLB) Act and as discussed in recent national publications. Next we compare class-size research to the SBR criteria as suggested by some authors and to other criteria for judging social programs. We summarize findings of important class-size studies and review theories that help explain the constant positive outcomes that should accompany class-size reductions (CSR), including a review of student behavior in small classes (Finn, Pannozzo, & Achilles, 2002 and submitted for publication).

We present a research-based summary of recommendations on how to implement small classes for maximum impact, and speculate about reasons that small classes are not, by now (after so many years of study) widely used in USA schooling. This section reviews cost issues and examples taken from actual sites, as differentiated from costs derived a) from models or estimates or b) from CSR that did not follow the research-based recommendations.

Although the rhetoric about the deplorable state of USA schools is strident, we believe that the schools that U. S. youth attend are reasonably good, given the overall situation. Nevertheless, most organizations can improve, and CSR improvements are likely to be efficient and effective if they build on SBR rather than on ideology and uninformed criticism.

An Impending Sense of Urgency

The topic of class size in elementary grades has taken on added importance recently. There are several reasons for the urgency. First, the compelling and long-standing research on class size is finally getting some attention. This attention may (should) escalate if educators hope to try to contend with the requirements of the federal No Child Left Behind (NCLB) Act, and especially the call for adequate yearly progress (AYP). Interestingly, a group of respected researchers pointed out some problems with the AYP goals and identified the high improbability that the goals can even be approximated (Linn, Baker & Betebenner, 2002)

Second, fiscal conditions throughout the nation (2001-2003, at least) are requiring state-level legislative and administrative actions to reduce deficits and looming debts. An option to balance budgets is often to increase class sizes and in other ways reduce education spending [e.g., delay capital outlay; cut "non-essentials" like art, music, drama (aren't these the basics? They surely came before reading, math, science!); reduce extra-curricular activities and athletics, curtail summer school]. This approach seeks to balance the budget by short changing students...
and requiring non-voting age youth to pay for adult ineptitude. However, by putting meaning into an inane motto, “Less is more,” it may be possible that fiscal hardships could benefit the class-size issue by urging educators both to be creative and to attend closely to the research on class size. Would people consider small classes, K-3, seriously if they knew that implementing them correctly would cost very little (if any) more once the space issue is resolved?

Third, the 2002 elections showed strong voter interest in education generally and in CSR specifically. Examples include election of “education” governors in Michigan, Wisconsin, Pennsylvania, Arizona, Kansas, Iowa, and Illinois, and a superintendent of public instruction in California, all of whom explicitly campaigned on attention to education, explicitly to class size. In Florida, voters passed Amendment #9 that directly addresses class sizes, especially in grades K-3. The victory margin of more than 200,000 votes was in the face of Governor Bush’s re-election and intense lobbing against class sizes. He was caught on tape saying he had “a couple of devious plans” if the amendment passed. (N. Y. Times, 11/1/02, p. A28 and Wall Street Journal 10/10/02, p. A4). He has started on a couple.

Interestingly, in August (2002) the Florida Association of District School Superintendents distributed a “White Paper” on “Class Size Reduction Constitutional Amendment” that generally opposed small classes for young students. One must presume that these educators did not clearly understand the research on class size because careful attention to the class-size research; cooperation among educators, the public, and policy leaders; and careful planning—even in the face of fiscal responsibility—can benefit students, help teachers and allow elected officials to make good on campaign issues.

To be useful, research must rest on a modicum of precision, clarity, and accuracy. Without accuracy, mandates for small classes may be translated into “devious” plans to employ some arithmetic substitution such as pupil-teacher ratio (PTR) or “average” class size for actual class size. Class size is a precise concept. Class size is determined by adding the number of students in each teacher’s class. Use of PTR (a ratio implies division) as a proxy for class size will produce PTR results—essentially minimal—instead of class-size outcomes. Thus, a proposed plan in Florida to compute the “average class size for each grade level by figuring averages for entire districts” (Editorial: Orlando Sentinel, 1/06/03) will assure at least a) continuation of non-class size outcomes in Florida and b) more “results” to show that small classes don’t matter, but once again using PTR rather than small classes as the improvement. A recent study has shown that in the USA the difference between class size and PTR in elementary grades is about n=10. (Sharp, 2002; Achilles & Sharp, 1998). So, if the PTR at a site is 16:1, the average teacher will face 26 or more students each day. This condition needs repair.

Two points are important. Following the research results closely will provide class-size outcomes similar to those found in the Student Teacher Achievement Ratio experiment (STAR), Wisconsin’s Student Achievement Guarantee in Education or SAGE demonstrations, in Burke
County, NC and in other places where class size has been implemented carefully in elementary grades. Recommendations for implementing appropriate-sized classes in elementary grades are in Attachment A, the last page in the paper. You are encouraged to duplicate and use it in your planning for class-size reduction.

**Some Background and SBR Support**

Class-size research has a long history. The designs and methods of class-size research have improved. Any discussion of implementing small classes can rely upon a research foundation that includes small and large studies and evaluations and on scientifically based research (SBR). As stated by Slavin (2002), the 2001 NCLB Act actually defined SBR as “rigorous, systematic and objective procedures to obtain valid knowledge.” This includes research that “is evaluated using experimental and quasi-experimental design” preferably with random assignment (p. 15, Emphasis added). Much class-size work (e.g., STAR, SAGE, evaluations of the Burke County, NC initiative) meets SBR tests such as experimental design, random assignment, longitudinal, and sufficiently large-scale. A summary of the SBR conditions of Project STAR is in Appendix A that describes STAR’s experimental design: SAGE and the other studies employed quasi-experimental designs.

The primary research base for class-size “impact” is Tennessee’s Project STAR (Word et al., 1990). As Mosteller (1995, p. 116) correctly noted, the Tennessee class-size effort really was three studies: the experiment in K-3 (STAR), checking on the continuation of benefits achieved in STAR (Lasting Benefits Study), and Project Challenge, a four-year study of class-size implementation. All of the STAR subsidiary studies and studies using the STAR database likely meet the new criteria for SBR. This SBR is available to guide CSR.

The NCLB points about SBR have been discussed in the *Educational Researcher* (Feuer, Towne, & Shavelson, 2002) where the authors summarized six criteria or principles of scientific endeavors (p. 7). Prior to contemplating the use of class-size research, consider how the substantial research base matches SBR criteria. One example showing how STAR meets the “Principles” of scientific endeavors (Feuer et al., 2002, p. 7) appears in Table 1. Table 2 compares STAR to Crane’s (1998) criteria for social-program research. An important point is that STAR and other class-size studies usually are independent research, a point made in Crane’s criterion 5. There is no monetary gain or connection (nothing for sale). Probably because of STAR’s simplicity, because no salesperson knocks on the door with a program to “save education,” and because of long-standing but erroneous claims that class size isn’t important, folks don’t really pay attention to the research. Besides, STAR results call into question much of current education practice. This idea is briefly discussed later.
Definitions are appropriate for research and for research papers. First, in this paper a “small” class (S) is about 14-17 students per teacher. A really small class would be tutoring. The definitions here appear in several sources by the authors of this paper.

**Class Size(s)** – “The number of students for whom a teacher is primarily responsible during a school year (Lewit & Baker, 1997, p. 113).” This is an addition problem. Class size is an organization for instruction important to teachers, parents, students.

**Pupil-Teacher Ratio (PTR)** – “The number of students in a school or district compared to the number of teaching professionals” (McRobbie et al., 1998, p. 4). In some venues all educators are part of the computation, including counselors, administrators, etc. In this division problem, the divisor is very important. PTR is a way to assure equitable distribution of funds and is important to administrators, policy persons, etc. The difference between PTR and class size in USA elementary schools is about n=10 (Sharp, 2002).

**Class-Size Reduction (CSR)** would include the processes involved in achieving class sizes smaller than the ones presently in place. Often this means changing the class size from 25 to 16 or so. One needs accurate pre and post data to support the change process.

**Average Class Size** is the sum of all students regularly in each teacher’s class divided by the actual number of regular teachers in those specific classes. If the 4 2nd grade rooms have 14, 16, 18, 18 (n=65) the average grade-2 class size is 16.25(or 16) students.

Data in large databases are PTR data. Surveys provide PTR data, or aggregated data (especially in secondary schools) of several classes resulting in estimated or average class size. Valid ways to get class-size data are 1) to count the students in a class and/or 2) to establish class sizes and then monitor them as in STAR. Class-size “research” cannot be done by substituting PTR numbers or outcomes for class size. Research on class size requires visits to schools to check actual class sizes! Appendix B summarizes differences between PTR and class size on five important dimensions that influence education practice and outcomes.

Tables 3 and 4 offer some of the extensive research base behind positive research outcomes listed in Tables 5 and 6. The large recent and sustained research base shown in Tables 3 and 4 provides a level of “confidence” in class-size outcomes. Table 5 describes (briefly) class-size benefits in 11 categories. Table 6 summarizes outcomes of small classes from a range of sources, including one review (i.e., Cooper) and some of the early studies and meta-analyses. Table 7 shows research and theories behind the class-size “impact,” including observed
processes and outcomes. Class-size outcomes are based upon more than 30 theories and established principles of effective schooling.

Cost Stories from Actual Sites.

Despite dire predictions of huge small-class costs that are based upon PTR models or upon PTR masquerading as class size, observations and analyses of actual small-class initiatives show that the goals can be achieved at little or no added costs if the implementation builds upon the research. Here are actual scenarios.

1. A small district in Michigan (n=600 pupils) achieved class sizes in K-4 of 15 or fewer pupils within the regular budget by re-assignment of personnel, some job consolidations, reduction of teacher assistants via attrition, and elimination of "projects" based on small-class benefits. Student achievement, teacher morale and district support are all demonstrably higher than before the CSR in this Michigan district. (e-mail personal communication)

2. A mid-size district (n=14,000 pupils) achieved class sizes of about 14 students, K-4, by careful planning, reductions in remediation projects as small-class outcomes allowed, by extending productive inclusion efforts, and by reallocating funds through attrition of teacher aides. After initial cost outlays for facility renovation, the small classes are within the normal state per-pupil expenditure range. Achievement in this previously low-achieving district is among the highest in North Carolina. The K-4 small-class benefits carry over into upper grades. (Egelson et al., 2002; Achilles, Harmon & Egelson, 1995).

3. As part of a major national study of class size and PTR from 1997-2002, Sharp (2002) conducted a case analysis of a single district's finances and resources to determine if the district could achieve small classes (n=14-16 pupils) in grades K-3. His results clearly affirmed that with careful planning, assessment of current program outcomes, and using only reallocation of current resources the district could achieve the goal of small classes for all K-3 students. The suggested reallocations were based on the expected outcomes of small classes as determined by the extensive class-size research.

Student Behavior and Safety

An obvious omission to date in discussions of class-size benefits and mechanisms for success has been The STUDENT. Rather, much class-size attention has focused on the design and strength of the studies, what teachers do or don't do, etc. Additionally, many class-size analyses have been directed to easily measured test-score advances. Although STAR researchers have explained research-based class-size benefits in four areas (see prior section on ABECEDARIAN concept), the public pressure for test score (scare?) gains and "quick fixes" has overtaken common sense about the purposes of education in the broader context.

There are compelling evidence and support for the idea that a powerful driver for small-class achievement in all ABCD areas is STUDENT BEHAVIOR or how students behave in various settings. (Finn, Pannozzo & Achilles, 2002 and in press). People behave differently in
different-sized groups. Research (e.g., Hall, 1966, 1976) has shown that space and crowding influence communication, learning, behavior, aggression.

The “Whys” of class-size successes are embedded in student engagement and its attendant attributes (Finn, 1993), the influences of space and crowding on that behavior, and what appropriate-sized classes allow students and teachers to do to improve learning (e.g., Darling-Hammond, 1998). The abstract from the Finn et al. paper summarizes student behavior as a pivotal reason for small-class success.

---

The “Whys” of Class Size: Student Behavior in Small Classes

Abstract

Small classes in the elementary grades boost students’ performance in all subject areas. However, researchers continue to seek a consistent, integrated explanation of “why” small classes have positive effects. This paper forwards the hypothesis that when class sizes are reduced, major changes occur in students’ engagement in the classroom. Engagement is comprised of “learning behavior” and pro-and antisocial behavior. Both are highly related to academic performance. The authors first review research on the relationship between class size and student engagement. Second, sociological and psychological theory about the behavior of individuals in groups is reviewed, to explain how student behavior can be affected by changes in class size. Both theory and empirical findings support our hypothesis, although additional research is required. In the conclusion, the authors propose that the same principles explain the effects of small schools and small learning communities on students’ attitudes and behavior.

---

The Overwhelming Conclusion from the Research

Changing the size of the class, the usual organization for delivery of instruction in U. S. schools, causes increased student outcomes (as shown in STAR). Unlike in a targeted project (e.g., Reading, where one would expect an increase in student reading scores), students in small classes improve in ALL subject areas tested (social studies, science, math, reading, spelling, etc.). But, not just test scores improve. Students improve in major ways that for ease in remembering, we have labeled: The ABECEDARIAN (ABCD) concept. The ABCD form is similar to Dr. James Comer’s four areas of schooling improvement, as well as results from the Perry Preschool experiment. (E.g., Schweinhart & Weikert, 1997; Weikert, 1998; Xiang & Schweinhart, 2002)

A Academics (e.g., test-score performance).
B Behavior and discipline in classes and in school, including safety.
C Citizenship and participation/engagement in and outside of school.
D Development into productive humane persons who contribute to society and are responsible for their actions.
Speculations About Reluctance to “Do” Small Classes

This section may ruffle some feathers. Good. Many people claim that educators are complacent and reluctant to change. Given the reluctance to operationalize what about 100 years of research has shown, we’d be inclined to agree. This section includes only an outline of ideas generated while this paper was in process and from observations of CSR.

Clearly, the barrier to implementing small classes in early grades is not parents or the citizenry in general. (What parents seek larger classes for their children?) What politician runs on a platform of larger classes? People really know the value of small classes. Elections in 2002 affirmed citizen support for the proven value of small classes. Voters elected class-size reform governors in many states: Arizona, Illinois, Iowa, Kansas, Michigan, Pennsylvania, and Wisconsin, a state with an effective class-size initiative. Voters passed a class-size constitutional amendment in Florida and overwhelmingly elected a class-size advocate as superintendent of public instruction in California. Small-class benefits thrive selectively. They should be available to all students, because they provide a base for excellent education, and all students are special. If the roadblock is not parents and citizens, then what is it? Our candidates are three groups of people and one tradition. The tradition began in 1965 and has not carefully been assessed in light of social changes and education demands. Here is an outline of key issues.

1. **Ideological differences.** Some groups strongly advocate their own ideas about education and constantly offer plans for education that have little or no support in education or social-science research. To advance their agendas, the groups attack small classes on a variety of fronts, often with no data or with hypothetical data: No space, no personnel, no money, no solid research, etc. A popular strategy is to parade pupil-teacher ratio (PTR) data that, truthfully, show little production gain and to claim that PTR equals class size. A second source of ideological differences relates to the value (both cost, and claims of effectiveness) of one improvement agenda over another. An example of this issue is the dissonance between “Teacher quality” (whatever that is), and class size (clearly defined and measurable).

2. **Administrator reluctance/inaction/lack of knowledge of the class-size research base.** The reluctance may be fed by the constant call among politicians and policy people for a “quick fix.” Small classes are primarily preventive, and not remedial. Thus, solid results of a small-class initiative will take 3-4 years to show up in the testing outcomes. Administrators need to be seen as “doing something—anything,” and small classes just seem too simple. Besides, if administrators keep plugging professional development for teachers, they can shift any blame for perceived school failure from themselves to teachers.

3. **Teachers** may vocalize support for small classes, but often when they realize that small classes may require other changes, they waffle (usually only briefly, until they begin to feel the professional reward of responsibility, success, and accountability). One superintendent of schools negotiated small classes, K-4, into the teacher contract. Teachers would have no more that 15 students per class (most had 13-14 students). In return, all funds were directed to teaching positions, reducing “support” personnel; teacher assistants were reduced via
attrition and those funds went into teacher salaries. At first, teachers complained that they had no place to send reluctant learners and discipline cases. (They were used to aides, specialists, transitional classes, etc.). The superintendent explained that they, the teachers, were the specialists and the “treatment” for the students was to be with them, the teachers, all day, each day as students learned what needed to be learned.

4. Tradition. Since about 1965, teachers have become accustomed to fragmentation in their lives, and to the disruption and lack of coherence in planning and instruction that accompanies sending difficult-to-teach kids to “specialists” (who are not accountable for the student’s basic test scores). One critique of an earlier “excellence movement” in education explained Title I and “project mentality” as follows (Wayson et al., 1988).

... reforms intensified the rigidities of the education system; they depersonalized the education process; they weakened the profession by creating splits between educators; they glorified specialization by elevating teachers to positions of dominance over other teachers; they narrowed roles for teachers; and they diminished power and respect for those who work most closely with children (p. 115).

Because many teachers today started teaching since Title I (1965) they are inured to this bleak scenario. The use of small classes in K-3 will get the system back on track—but the teachers must step up to be counted, for class-size success will require them to assume a professional case-load responsibility of a workable number of students. (This does NOT mean the elimination of all specialists and projects. It does mean a reassessment of business as usual, using what we know now and accommodating diversity, ESL, inclusion . . . ).

The difference between class size and PTR in U. S. schools (about n=10) provides one place to start planning how to achieve small classes. Small-class benefits such as reduced grade retention, and need for remediations, and successful inclusion outcomes provide short-term cost relief. Because appropriate-size classes impact students and teachers (and parents, too, in some cases), small classes are an incentive to attract and keep teachers, a potential factor in any teacher shortage.

Careful attention to class-size research could produce small classes at reasonable costs by eliminating ideas never shown effective in SBR as called for in the new NCLB legislation. Let’s start the discussion.

Table 7 provides an outline for understanding small classes in five general areas: 1) learning, 2) teaching, 3) classroom, 4) “other,” and 5) student behavior. These points serve as guides for discussion and planning. Rather than a panoply of “projects” to get each one of these desirable education interventions, contexts, or outcomes, reducing class sizes to fit the important task at hand causes, or paves the way for, each of the elements to “impact” the education enterprise. What other intervention is so comprehensive? Small classes are “whole school reform,” clearly meeting the idea of “scientifically based research.” The students deserve better.
TABLE 1. COMPARISON OF STAR DESIGN, PROCESSES, AND FACTS WITH ONE SET OF “PRINCIPLES OF INQUIRY”.

Although no universally accepted description of the principles of inquiry exists, we argue nonetheless that all scientific endeavors:

<table>
<thead>
<tr>
<th>SCIENTIFIC ENDEAVORS*</th>
<th>STAR DESIGN, PROCESSES and FACTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <strong>Pose Significant Questions That Can Be Investigated Empirically.</strong></td>
<td>1. STAR was driven by two significant, major questions: What is the EFFECT of small classes in primary grades on the 1) Achievement and 2) Development of students? Researchers addressed secondary questions required or implied in the legislation: Effects of a) full-time teacher aide, b) training, c) duration, d) cohort, e) random assignment. (See Table 5). Researchers studied other questions: teacher quality (by credentials), comparisons of sample with state averages, checks on “randomness,” time use, teaching processes, incentive value . . .</td>
</tr>
<tr>
<td>The initiating law required questions and processes. Researchers added others</td>
<td>2. STAR was deeply rooted in prior research and theory. Theories are evident in the design, data forms, analysis steps. Additional theory and refinements were “teased out” during the study (1984-1990), as data were analyzed (some data still await analysis), as STAR played into Project Challenge, and while students progressed throughout their schooling for longitudinal results (they would graduate from High School in 1998, if on schedule).</td>
</tr>
<tr>
<td>2. <strong>Link Research to Relevant Theory.</strong></td>
<td>3. “Effect” required an EXPERIMENT (Campbell &amp; Stanley, Design #6), of sufficient Duration (4-years), Magnitude (at least 80 classes of each type – eventually 11,600 students). The experimental plan was small class (S) at 13-17; regular (R) at 22-25; and full-time Aide (RA) at 22-25. Within-school design was parsimonious, reduced school-level effects, eliminated control group mortality, moderated the “Hawthorne Effect” if it might be a factor (Appendix A summarizes the experiment).</td>
</tr>
<tr>
<td>(STAR began in 1984, so some design and theory issues we now know (2003) were not yet refined. Table 8 is a summary of some theories supporting STAR.</td>
<td></td>
</tr>
<tr>
<td>3. <strong>Use Methods That Permit Direct Investigation of the Questions.</strong></td>
<td></td>
</tr>
<tr>
<td>The variable of focus was class size so only class size was manipulated; the Aide was a Pupil-Teacher Ratio (PTR) element. STAR represented school as it is normally operated.</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 1. COMPARISON OF STAR DESIGN, PROCESSES, AND FACTS WITH ONE SET OF “PRINCIPLES OF INQUIRY”** (con’t)

“ALTHOUGH NO UNIVERSALLY ACCEPTED DESCRIPTION OF THE PRINCIPLES OF INQUIRY EXISTS, WE ARGUE NONETHELESS THAT ALL SCIENTIFIC ENDEAVORS: …

4. **Provide a Coherent and Explicit Chain of Reasoning.**

   Longitudinal class-size studies were needed to test duration. Without an experiment, effects of SES, teacher, principal leadership (etc.) clouded the class-size issue/effects.

5. **Yield Findings That Replicate and Generalize Across Studies, and:**

   Work continues here as more states, and local districts move into class-size changes. Note International work in Australia, England, Netherlands, Sweden. (see also Tables 3 and 4)

6. **Disclose Research Data and Methods To Enable and Encourage Professional Scrutiny and Critique.***

   STAR data, methods and outcomes are in the Final Report, papers and articles by the PIs, dissertations, and other print sources.

4. Much of the reasoning appears in the STAR Report literature review, data instruments, observation data, research questions, sample, and design. Prior to STAR there was disagreement on the effects of group (class) size on student outcomes. Before establishing statewide class-size limits, Tennessee lawmakers and policy persons sought evidence about class size and paraprofessionals. They commissioned STAR

5. STAR results have been replicated and generalized in state studies (e.g. SAGE in WI); by state law (e.g. HB 72 in TX); in observations (SSS); in cases studies (e.g. Rockingham Co, NC); in large (n=15,000) and small (n=1200) districts (Burke, Co. NC; Litchfield, MI); in Title I schools (n=16) in a large district; in single schools (SC, NC, LA). “Micro” comparisons contrast with “macro” or statewide events (e.g., NC, TN, TX, IA, UT) and even in NV that did some PTR and in CA, a “near text-book case of doing it wrong” (Biddle & Berliner, 2002). Results are always positive.

6. The Spencer Foundation assisted PI’s to organize, clean, and post STAR data on “The Web.” After the final report was accepted, data were provided to researchers in London and later to persons in the USA. Critique is evident in some journal articles. “Scrutiny” is in the hands of the secondary analyzers, and has seldom been rigorous, absent pre-conceived ideology.

---


*** The narrowness of most STAR critiques suggests that the STAR Report and Papers (The Primary Sources) were read by few (e.g., Mosteller, 1995; Burke, Co. administrators; SAGE staffers; SERVE personnel, Doctoral Students); Few persons engaged the four Principal Investigators (PI’s) in discussions or asked important questions so they could understand STAR outcomes. Professor Mosteller (1995) actually explained that in reality STAR was THREE studies. (STAR, LBS, CHALLENGE)
Table 2. Critique of STAR Results Using Crane's (1998) Criteria. STAR is a Class-size Reduction (CSR) Experiment, Not a Pupil-teacher Ratio (PTR) Effort.*

<table>
<thead>
<tr>
<th>CRANE CRITERIA and QUESTIONS</th>
<th>STAR'S FACTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Do the benefits outweigh the costs? <strong>YES</strong></td>
<td>1. In the short term (K-3), there were no definitive data. In the &quot;follow-up studies,&quot; <strong>yes</strong>; in the STAR reanalysis, <strong>yes</strong>; in alternative implementations, <strong>yes</strong>. See Krueger (1999; Finn &amp; Achilles, 1999; Finn et al., 2001).</td>
</tr>
<tr>
<td>2. Does the program have a statistically significant effect on the treatment group? <strong>YES</strong></td>
<td>2. Yes. This statistically significant difference was found each year, all years, and in many combinations of analyses done by STAR persons and by others (as far away as London).</td>
</tr>
<tr>
<td>3. What is the magnitude of the program's effect? <em>(Shown in Effect Size or ES)</em></td>
<td>3. Effect-size (ES) results were .17-.40 in the early analyses. Effects were about twice as high for minority children as for Anglo children, grades K-3 (each year, all years). Grade-equivalent analyses show continuing growth even after students leave small classes (see #4). (Finn &amp; Achilles, 1999; Finn et al., 2001).</td>
</tr>
<tr>
<td>4. How long do the effects of the program last? <em>(At least into high school and beyond.)</em></td>
<td>4. Positive academic and social effects of K-3 small classes are highly visible in H.S. and beyond including in college-entrance tests. (Boyd-Zaharias &amp; Pate-Bain, 2000; Krueger, 1999; Krueger &amp; Whitmore, 2000).</td>
</tr>
<tr>
<td>5. What is the relationship of the evaluator to the program? <em>(Independent)</em></td>
<td>5. The STAR evaluator was a contracted independent expert. STAR personnel did secondary analyses. The external expert’s work is (and was) the primary analysis accepted and published. Others have re-analyzed STAR data with similar results.</td>
</tr>
<tr>
<td>6. Can the program and its results be replicated? <em>(Yes)</em></td>
<td>6. &amp; 7. They have been consistently replicated in well designed class-size analyses. Replications of STAR have been achieved in single districts, and in general policy implementations. Reported gains and ES for well conducted studies are similar. Evaluations of state-wide small-class efforts in CA, and the results in Texas (HB 72, 1984) suggest large-scale benefits, but these results are less definitive than STAR or SAGE in Wisconsin, probably because of less controlled implementations.</td>
</tr>
<tr>
<td>7. Can the program maintain its effectiveness on a larger scale? <em>(Still being assessed. Yes if well implemented.)</em></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Summary Listing of Some Class-Size Studies and Research Summaries, 1970-2002: Thirty (+) years of The "Present Generation."

<table>
<thead>
<tr>
<th>Author, Study</th>
<th>Source/Date *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lindbloom</td>
<td>1970</td>
</tr>
<tr>
<td>Olson</td>
<td>1971 (From Cavenaugh, 1994)</td>
</tr>
<tr>
<td>Glass &amp; Smith</td>
<td>1978, 1979</td>
</tr>
<tr>
<td>Smith &amp; Glass</td>
<td>1979</td>
</tr>
<tr>
<td>Filby et al.</td>
<td>1980</td>
</tr>
<tr>
<td>Glass et al.</td>
<td>1982</td>
</tr>
<tr>
<td>Shapson et al.</td>
<td>1980</td>
</tr>
<tr>
<td>Evertson &amp; Folger</td>
<td>1989</td>
</tr>
<tr>
<td>Evertson &amp; Randolph</td>
<td>1989</td>
</tr>
<tr>
<td>STAR (Generally)</td>
<td>Word et al. (1990); Johnston (1990)</td>
</tr>
<tr>
<td>Teacher Interviews (1000+)</td>
<td>Bain et al. (1992)</td>
</tr>
<tr>
<td>Robinson</td>
<td>1990 Research Review</td>
</tr>
<tr>
<td>STAR Good Teacher Study</td>
<td>Bain &amp; Lintz</td>
</tr>
<tr>
<td>Project Success (NC)</td>
<td>1994 (In Achilles et al., 1994)</td>
</tr>
<tr>
<td>Wenglinsky</td>
<td>1997 (ETS)</td>
</tr>
<tr>
<td>Participation &amp; Engagement</td>
<td>Finn (1998, 1993), Voelkl</td>
</tr>
<tr>
<td>California CSR (STAR-Related)</td>
<td>CSR Consortium (1999), Bohrnstedt, etc.</td>
</tr>
<tr>
<td>Long-Term Effects (STAR)</td>
<td>1999 (  )</td>
</tr>
<tr>
<td>Teacher Aides</td>
<td>Krueger, Bain et al.</td>
</tr>
<tr>
<td>College Entrance Tests</td>
<td>Finn et al., (2001), Nye et al. (2002)</td>
</tr>
<tr>
<td></td>
<td>Finn, Gerber et al. (2001)</td>
</tr>
<tr>
<td></td>
<td>Bain, Boyd-Zaharias, Achilles</td>
</tr>
<tr>
<td></td>
<td>Krueger &amp; Whitmore (2000, 2002)</td>
</tr>
</tbody>
</table>

* Many of these studies have been reviewed briefly in Achilles (1999) Let's Put Kids First. The work of B. Bloom on tutoring and the "2-Sigma Problem" is foundational.
Table 4. Samples of the STAR Legacy of Class-size Studies, Categorized as “Subsidiary” (directly from STAR), “Ancillary” (using the STAR database) and “Related” (usually involving STAR researchers and expanding STAR earlier findings).

<table>
<thead>
<tr>
<th>CATEGORY, TITLE &amp; PURPOSE *</th>
<th>DATE(S)</th>
<th>AUTHOR(S), SOURCE, DATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAR Pilot (DuPont)</td>
<td>1984-1986</td>
<td>Bain et al. 1984, 1985</td>
</tr>
<tr>
<td>Subsidiary Studies</td>
<td></td>
<td>Finn &amp; Achilles, 1990</td>
</tr>
<tr>
<td>- Participation, Grades 4, 8</td>
<td>1990, 1996</td>
<td>Finn, 1989, 1993; Voelkl, 1995; Finn et al., 1989; Finn &amp; Cox, 1992</td>
</tr>
<tr>
<td>Ancillary Studies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Re-analysis, Sample “drift” (out-of-range classes)</td>
<td>1985-2001</td>
<td>Boyd-Zaharias et al., 1995</td>
</tr>
<tr>
<td>- Teacher Aides</td>
<td>1990-2002</td>
<td>Achilles et al., 1994; Finn et al., 2001; Boyd-Zaharias &amp; Pate-Bain, 1998</td>
</tr>
<tr>
<td>- College entrance exams</td>
<td>1999-2001</td>
<td>Finn, Achilles et al.; Bain et al.</td>
</tr>
<tr>
<td>Related Studies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- SERVE Studies in NC</td>
<td>1994-2003</td>
<td>Achilles et al., 1995; SERVE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Krueger, 1997-2002</td>
</tr>
</tbody>
</table>

* This is a sample of STAR-related class-size studies. Not all authors appear exactly as listed here. A similar table appears other STAR reports and articles. Several have reported on STAR (e.g., Mosteller, 1995; Finn, 1998; McRobbie et al., 1998); several non-STaR persons have conducted secondary re-analyses of STAR data (e.g., Goldstein & Blatchford, 1998; Krueger, 1997-2001, Nye et al., 1999-2002). Many single-district studies, dissertations and re-analyses.
Table 5. Synopsis of Class-Size Findings from STAR and Various Other Sources.*

<table>
<thead>
<tr>
<th>Findings, Idea, Issue</th>
<th>Selected Sources of Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Class-size effect was found in all sites, for all participants, at all times and grades K-3. This includes tutoring and &quot;special&quot; projects.</td>
<td>STAR, Challenge, Reading Recovery (RR); Success for All (SFA)</td>
</tr>
<tr>
<td>II. Small classes work best when students start (K, 1) school in them; they are preventive, not remedial. Formal and small-class education MUST start no later than K, be intense (all day, every day) and last at least 3 years (Duration).</td>
<td>STAR, SSS, Challenge SAGE, Burke Co., Abecedarian (NC), Finn &amp; Achilles (1999) Perry Pre-School, Finn et al., 2001</td>
</tr>
<tr>
<td>III. Crowding, not just small classes, is an issue. School safety and environment are improved. (Prout, 2000). School size is important.</td>
<td>STAR, SSS, K. Nye, Fowler &amp; Walberg, Behavioral Research, Cotton, others.</td>
</tr>
<tr>
<td>IV. Although all pupils benefit from small (S) classes in K-3, some students benefit more than others.</td>
<td>STAR, SFA, RR, LBS, Other class-size work. Robinson (1990).</td>
</tr>
<tr>
<td>V. The teacher is important. Each pupil's learning depends upon the teacher and others in the class. (Thus the class is the unit of analysis).</td>
<td>STAR, LBS, SSS, Challenge, Burke County, CSR in California.</td>
</tr>
<tr>
<td>VI. A teacher aide does not improve student outcomes. This adds to crowdedness and causes new dynamics (Issues: Training, inclusion, ESL, role description).</td>
<td>STAR, Other Studies. Finn, Gerber et al., (2001); Bain &amp; Boyd-Zaharias (1998); Gerber et al. (2001).</td>
</tr>
<tr>
<td>VII. Teachers should use known educational-improvement processes: (Parent and home involvement, portfolios, alternative assessments, etc.). Small classes may not change what teachers do—just how much they do good things well.</td>
<td>STAR, LBS, SSS, Filby et al., Burke County, NC; Downtown School, NC STAR Teacher Studies. Achilles, 1999, 2002, 2003</td>
</tr>
<tr>
<td>VIII. Reduce retention in grade especially when student will be moving into another small class. (Retention should not be used, unless in extreme cases).</td>
<td>STAR, Many studies of Retention (Holmes and Matthews).</td>
</tr>
<tr>
<td>IX. Study costs and benefits; Use PTR and class size differences to get to small classes.</td>
<td>STAR, SSS, PTR studies, Sharp, Darling-Hammond; Miles</td>
</tr>
<tr>
<td>X. Small classes and small schools encourage increased student participation in schooling. (Engagement)</td>
<td>Finn, Voelkl, STAR, LBS, Lindsay's work, etc., Finn et al., (2001; 2002)</td>
</tr>
</tbody>
</table>

* Detailed references are available. They were omitted because of space. RR = Reading Recovery; SFA = Success for All; SSS = Success Starts Small.
Table 6. Summary of Small-Class Benefits: Source, Study, (x = Yes, as included in the source named)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Increases:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Time on Task</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>• Hands-on</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>• Indiv. Attn.</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>• Diagnosis</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>• Social Climate</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>• Management</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>• Participation</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>• Academics</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>• Parent Involv.</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>• Early ID of Spec. Ed.</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>• Morale</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>• Space</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>• Enrichment</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>• Text/Methods</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>• Group Work</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>8</td>
<td></td>
</tr>
<tr>
<td><strong>B. Decreases:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Indiscipline</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>• Retention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>• Spec. Ed.</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>• Stress</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

Table 7. Small Class (K-3) Benefits Are Supported by Research and Established Theories About Groups, Teaching, Learning, and School Outcomes.

I. LEARNING
A. Task Induction: Learn About School (Student’s Work).
B. Participation, Engagement, Identification.
C. Mastery of Basics
D. Time On Task Increases.
E. Appropriate Homework
F. Child Development/Developmentally Appropriate.
G. Early Intervention, Duration

II. TEACHING
A. Individual Accommodation.
B. Early Diagnosis And Remediation Of Learning Difficulties.
C. Teach To Mastery.
D. Immediate Reinforcement.
E. Assessment (In-Class)
F. Portfolios, Running Records.
G. Effective Teaching Methods.
H. Planned, Coherent Lessons (Seamless Transitions).

III. CLASSROOM
A. Classroom Environment (E.g.: Air Quality, Materials, Space, Crowding, Noise).
B. Personal Attention/Community.
C. Inclusion, Special Needs
D. Variable Room Arrangements (E.g., Learning Centers).
E. Classroom Management.
F. Less Indiscipline
G. Many Volunteers.

IV. “OTHER”
A. Increased Parent Interest.
B. Reduced Grade Retention
C. Increased Teacher/Student Morale/Energy.
D. Teacher Accountability and Responsibility
E. Few Projects and “Pull Outs.” (Coherence). Intensity
F. Student-Led Activities
G. Assessment (Outcome)
H. Field Trips Possible with Fewer Adults/Smaller Vehicles

V. STUDENT BEHAVIOR (B)*
1. Class size and Engagement: More Engaged in Learning and Pro-social (B) and Less in Disruptive (B). Principles: 1) “Visibility of the Individual” a) Time per Student, b) Diffusion of Responsibility and c) Social Loafing; 2. Sense of Belonging a) Group Norms [e.g., Learning (B)] Influences All Members, b) Psychological Sense of Community (PSOC) such as Support and Inclusion. Results are Similar to School Size Work.

Appendix A

A Longitudinal Class-Size Experiment: Scientifically Based Research.

STAR (1985-1989) and the many studies that build upon STAR benefit from the experiment's tightly controlled, in-school, randomized longitudinal design. STAR was conducted by a four-university consortium with considerable external support from consultants, advisory groups, and the Tennessee State Department of Education. Basic design issues are:

(1) STAR was a controlled four-year longitudinal experiment that permitted, to the extent possible with empirical data, causal conclusions about outcomes. Pupils entering K were randomly assigned to a small class (S; 13-17), a regular class (R; 22-27), or a regular class with a full-time teacher aide (RA). Pupils entering in later years were assigned at random to classes. Teachers were assigned at random. Randomization and testing were monitored carefully.

(2) Built on prior research STAR began in primary grades. Small classes had fewer than 20 students. STAR’s post-test only design. (Campbell & Stanley, 1963) enabled researchers to study the effects on minority, majority, male and female students. The design produced a "real" difference in the class sizes, from an average of 24 pupils to an average of 15.

(3) The samples were large and diverse. The K year involved over 6300 students in 329 classrooms in 79 schools in 46 districts. The first-grade sample was larger still. The large samples were maintained throughout the four years, producing an excellent longitudinal database. Total sample = 11,601.

(4) With minor exceptions, students were kept in their class in grades K-3 (cohorts). A new grade-appropriate teacher was assigned to the class each year.

(5) The class arrangement was maintained throughout the day, all year long. There was no intervention other than class size and teacher aides. Teachers received no special training except for a small sample in second grade; no special curricula or materials were introduced. (Training didn’t increase outcomes).

(6) Norm-referenced tests (NRT), and criterion-referenced tests (CRT) and measures of self concept and motivation were administered each spring. Students were aggregated to classes and classes nested into schools for analyses. Teachers and teaching were studied, as were grade retention, participation, aide use, etc.

(7) Students were followed and evaluated after STAR ended in grade 3. Most students graduated in 1998. Their college-entrance test results were monitored. (Krueger & Whitmore, 2000). Dropout rates were analyzed Pate-Bain, Boyd-Zaharias, Finn, 2003.
Appendix B

Some Major Differences Between Class Size (CS) or Class-size Reduction (CSR) and Pupil-Teacher Ratio (PTR).

<table>
<thead>
<tr>
<th>VARIABLES of note in comparing PTR and CS</th>
<th>PUPIL-TEACHER RATIO (PTR)</th>
<th>CLASS SIZE (CS) or (CSR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition</td>
<td>Students (n) at a site (building, district, class) divided by: teachers, educators, adults, (etc.) serving the site.</td>
<td>Students (n) in a teacher’s room regularly, and for whom the teacher is accountable.</td>
</tr>
<tr>
<td>Computation</td>
<td>DIVISION, with various divisors available depending upon the EXACT definition.</td>
<td>ADDITION. This cannot be accurately determined from large databases.</td>
</tr>
<tr>
<td>Concept</td>
<td>The teacher needs help; the student needs special services the teacher cannot provide.</td>
<td>A competent teacher can handle most education issues if given a reasonable case load.</td>
</tr>
<tr>
<td>Operation and Context</td>
<td>A project and “pull-out”- driven model full of commotion and “Band Aid” treatments. Loss of time on task. Difficulty in determining responsibility and accountability.</td>
<td>Teacher is responsible and accountable for the student’s growth and development: Academics, Behavior, Citizenship, Development, (A, B, C, D) Small focused learning groups.</td>
</tr>
<tr>
<td>Outcomes</td>
<td>CONSISTENTLY MARGINAL. Education “production function” analyses (Hanushek, 1998); Boozer and Rouse (1995); Title I evaluations, Borman and D’Agostino (1996) Wong and Meyer (1998), etc.</td>
<td>CONSISTENTLY POSITIVE on many variables (A, B, C, D). See class-size results from many studies. There is much consensual validation, anecdotal evidence, and “common-sense” support.</td>
</tr>
</tbody>
</table>
Recent examples of the PTR and class-size confusion (e.g., just add teachers) and mis-use of the terms in articles and policy pieces include:


- Several policy papers from “Think Tanks” such as The Heritage Foundation, e.g.:

  Johnson, K. A. (6/9/00). Do Small Classes Influence Academic Achievement? What the National Assessment of Educational Progress Shows.
References and Bibliography


REF-3


SERVE. See under author names: e.g., Egelson


Authors’ Notes

Comments and opinions expressed in this paper are those of the author; they do not necessarily reflect the positions and policies of supporting agencies, or of any other person.

1. C. M. Achilles, currently a Professor Education Administration at Eastern Michigan University and Seton Hall University (both part-time) was one of four Principal Investigators (PIs) of STAR and a consultant on numerous class-size studies (1984-present), including PI of Success Starts Small (SSS), a year-long observation study in matched schools but with different class sizes, grades K-3. Jeremy D. Finn, Professor in the Graduate School of Education, SUNY Buffalo was the design and analysis consultant for Project STAR. He has continued to analyze STAR and STAR-generated data (1985-present). He consults on class-size issues and evaluated the Buffalo, NY class-size effort. Work on the Student Behavior theory was performed with the support of a grant from the Spencer Foundation, “Class Size and At-Risk Students.”

This paper provides support for remarks made at the group session symposium. It is not intended to contain the actual remarks, but provides detailed background information, history, and references related to class-size issues.

2. Much of the text, many ideas and most research results contained in this paper have been presented in prior papers by the same authors. The papers are cited then listed in a long references section that doubles as a class-size bibliography. One paper used extensively here was developed for the SouthEast Regional Vision for Education (SERVE) as a handout [Achilles, C. M. (2003), “The impact of class-size reduction” (CSR)]. Another paper is in review for publication [Finn, J. D., Pannozzo, G. M., & Achilles, C. M. (2002). “The ‘whys’ of class size: Student behavior in small classes”].

The authors thank research-and-practice comrades who have been and who are working to achieve appropriate-sized classes for the difficult task of schools. A few of these persons are Helen Pate-Bain, Paula Egelson, Pat Harman, Art Hood, Jayne Boyd-Zaharias, Sheldon Etheridge, Gilda Howard-Outz, Mark Sharp, persons in Burke County and Rockingham County, NC, the many local school administrators who strive to improve teaching conditions so teachers can teach well, legislators and policy persons who seek ways to improve class conditions so students can learn.

3. Although we hope that the ideas presented here will generate thought and discussion—even be a bit contentious—nothing here should be seen as disparaging prior (or future) class-size work, critique, or critics. We believe that the ideas need to be made public for comment.


The similarity of the Comer efforts and of the four points in the “ABECEDARIAN Compact” for small-class outcomes helps emphasize that class size is a concept rather than a “program” and that by adjusting class size we can anticipate an array of important schooling outcomes, not just improved academic achievement.

The Perry Preschool Program, another randomized education experiment that has followed its subjects from preschool into adulthood, shows that early intervention and small classes provide positive, measurable, short and long-term cognitive (academic) and non-cognitive (social) benefits.

5. The following is an outline of a plan to get small classes, K-4, at no added costs in a small school system of about 600 students.
Due to declining enrollment no budgetary controls our district was facing severe budget issues. A restructuring committee had been formed and we had already laid off 7 teachers. Needless to say, staff morale around here the last couple of months has been horrible. Equally important, our layoffs had led to large elementary classes -- with aides.

When I heard your talk the nagging doubts I had exploded into the reality that we were doing everything WRONG for kids. However, I also knew that we could NOT afford anything different. THEN, you had the audacity to say that as superintendents we could control things and make this happen. I knew right then that you didn't know what it took to be a public school superintendent in the new millennium. BUT, because what we were doing was wrong for kids, I decided to step outside the box and give it a try.

IT WILL WORK!! There are a few "bugs" to work out with the unions (teachers & support staff) but they are excited for the first time in months. The staff feels like instead of curling up in the fetal position while the bear attacks -- hoping it only mauls us badly instead of killing us -- WE ARE FIGHTING BACK!

Here's what will happen:
- class size capped at a maximum of 15 students (will be written into contract) for grades k-4!!
- there will be no classroom aides, except as required by special education
- there will be no elementary principal, just one k-12 principal
- there will be no elementary counselor, just one k-12 counselor
- to promote respect/responsibility/teamwork, the teachers will supervise the students in the cleaning of their classroom at the end of the day (We can reduce some custodial positions and focus the remaining personnel on the public/common/grounds areas)
- elementary teachers will do their own specials (art, music, PE)
- however, teachers may combine classes for specials only -- not academic classes
- this will give the teachers some planning time
- teachers who used to teach art, music, PE will return to the classroom
- elementary teachers will supervise their own students for one of two recesses
- saves some money for playground supervision

6. This fragmentation may have been one way to address large classes and the growing diversity in schools caused both by desegregation and immigration. Designation as "specialists" or as a "project director" was a way to reward good teachers—by taking them out of classrooms! That was then. Now is now. Title I was the chief architect of this way of working and consistent evaluations have shown that Title I isn't particularly effective [e.g., Abt, 1997; Borman & D'Agostino, 1996; Wong & Meyer, 1998]. It is time to adjust Title I in accordance with more than 100 years of substantive research.
ATTACHMENT A

RECOMMENDATIONS FOR CLASS-SIZE CHANGE

As the move to implement appropriate-sized classes in America’s public schools escalates, educators have much information available. From years of studying and observing small classes, researchers and scholar practitioners have compiled a research base, theories, and exemplary practices of outstanding teachers to guide effective implementations of small classes. Informed Professional Judgement or IPJ is at the heart of class-size changes. **SMALL CLASSES ARE NOT SIMPLY HIRING TEACHERS AND DOING BUSINESS AS USUAL.** A class-size initiative will incorporate what the long-term class-size research has determined are important steps for successful schooling outcomes.

1. **EARLY INTERVENTION.** Start when the pupil enters “schooling” in K or even pre-K.

2. **SUFFICIENT DURATION.** Maintain the small-class environment for at least 3, preferably 4, years for enduring effects. Encourage parent involvement in schooling.

3. **INTENSE TREATMENT.** The pupil spends all day, every day in the small class. Avoid Pupil-Teacher Ratio (PTR) events, such as “pull-out” projects or team teaching. Develop a sense of “community,” close student-teacher relations, and coherence.

4. **USE RANDOM ASSIGNMENT** in early grades to facilitate peer tutoring, problem-solving groups, student-to-student cooperation, and active participation and engagement. (STAR).

5. **EMPLOY A COHORT MODEL** for several years so students develop a sense of family or community. STAR results show the power of both random assignment and a cohort model. “Looping” adds teacher continuity to the cohort, and may be a useful strategy for added benefits. (Research is needed here).

6. **EVALUATE** process and outcomes carefully, and share results. Appropriate-sized classes in elementary grades will take policy and perhaps even legislation change.

The difference between the PTR and actual class size provides some guidelines for planning. If the site has a PTR of 12:1, that suggests enough personnel to work toward class sizes of 15:1 or so and still keep some teachers for special assignments.

Adding ever endless “projects” ala Title I and continually disrupting the teacher’s and students’ day and continuity (e.g., coherence and stability) are not what the class-size research is about. To avoid needless costs and confusion, start in K and 1, add a grade per year through third grade. Reduce “specials” as small-class benefits will allow and re-allocate personnel to teach small classes.

---

1 This compilation of class-size information from many studies and from practice appears in similar form in several papers by C. M. Achilles and J. D. Finn.
I. DOCUMENT IDENTIFICATION:

Title: School Improvement Should Rely on Reliable, Scientific Evidence. Why Did "No Child Left Behind" Leave Class Size Behind? (Symposium) (But contains only paper)

Scientifically-Based Research (SBR) ... Shows How to Leave No Child Behind...

Authors: C.M. Achilles J.D. Finn

Corporate Source: Paper at Am Assn of School Administrators

Publication Date: 2/23/03

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced to the monthly abstract Journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

... __________________________

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 1

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROCIDE, AND IN ELECTRONIC MEDIA, FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

... __________________________

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 2A

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROCIDE ONLY HAS BEEN GRANTED BY

... __________________________

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 2B

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only.

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only.

Documents will be processed as indicated provided reproduction quality permits.

If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, please

[Signature]

[Printed Name/Position/Title]

[Professor Ed Admin]

[Organization/Address]:

[Achilles] 4477 Snug Harbor

Geneva, NY 14456

[Telephone]:

[215-956-3695]

[FAX]:

[Email Address]:

[Date]:

[3.20.03]

E-mail: achilles@rockwell-intl.com"
I. DOCUMENT IDENTIFICATION:

Title: School Improvement Should Rely on Reliable, Scientific Evidence. Why Did “No child left Behind” Leave Class Size Behind? (Symposium) (But contains one paper)

Scientifically-Based Research (SBR)... shows how to leave no child behind...

Author(s): C.M. Achilles J.D. Fink

Corporate Source: Paper at Am Assn of School Administrators

Publication Date: 2/23/03

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 1

[ ]

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 2A

[ ]

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only.

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 2B

[ ]

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only.

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature: C.M. Achilles

Printed Name/Position/Title: C.M. Achilles

Organization/Address: 4777 NUG HARBOR

Telephone: 315 967 2395

E-Mail Address: plato936@rochesters.rr.com

Date:
III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:

Address:

Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

N/A

Name:

Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC Clearinghouse on Educational Management
1787 Agate Street
5207 University of Oregon
Eugene, OR 97403-5207

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2nd Floor
Laurel, Maryland 20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

3 (Rev. 9/97)
US VERSIONS OF THIS FORM ARE OBSOLETE.