

The “Why’s” of Class Size: Student Behavior in Small Classes

Jeremy D. Finn and Gina M. Pannozzo

University at Buffalo, State University of New York

Charles M. Achilles

Eastern Michigan University

Small classes in the elementary grades have been shown to boost students' academic performance. However, researchers continue to seek a consistent, integrated explanation of “why” small classes have positive effects. This article forwards the hypothesis that when class sizes are reduced, major changes occur in students' engagement in the classroom. Engagement is composed of “learning behavior” and pro- and antisocial behavior. Both are highly related to academic performance. We first review research on the relationship between class size and student engagement. Second, we review sociological and psychological theory about the behavior of individuals in groups 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. High-priority questions needing further research are identified in the conclusion.

KEYWORDS: class behavior, class size, engagement, small classes, student behavior, teacher-student interactions.

Researchers have studied class size in American schools for more than a century. Well over 100 studies were reviewed by Glass and Smith (1978) and Robinson (1990). These authors concluded that the preponderance of the evidence showed that reduced-size classes—below 20 pupils—were associated with improved academic performance. Effects were most pronounced in the early primary grades and especially among students from low-income homes. All three findings were substantiated in Tennessee's Project STAR (Finn, Gerber, Achilles, & Boyd-Zaharias, 2001; Word et al., 1990). STAR differed from previous research because it was a large-scale randomized experiment having both short- and long-term follow-ups. Its findings were confirmed by several research teams using different statistical approaches, and the results have been replicated in other evaluations, including Tennessee's Project Challenge (Achilles, Nye, & Zaharias, 1995) and Wisconsin's Project SAGE (Molnar, Smith, & Zahorik, 2000). The success of these initiatives motivated many states and districts to undertake class-size reduction programs.

Although the findings are strong and consistent, they have met with some skepticism. In particular, economist Eric Hanushek has disputed whether small classes

are efficacious. Based on his own review of several hundred studies, he stated: “Extensive econometric investigation show [*sic*] NO relationship between class size and student performance” (1998, p. iii). Debate about this proposition has been heated, and several researchers even reanalyzed Hanushek’s own data, reaching the opposite conclusion (e.g., Hedges, Laine, & Greenwald, 1994; Krueger, 1999). More importantly for the present review, the studies summarized by Hanushek were not studies of class size¹ but of a different construct: the pupil-teacher ratio of schools, districts, states, and countries. Aggregate pupil-teacher ratios do not describe the day-to-day setting in which students are learning; many districts have low pupil-teacher ratios, while most students spend the entire school day, every day, in crowded classrooms (Finn & Achilles, 1999; Lewit & Baker, 1997; Miles, 1995).

This review focuses on the size of individual classes and addresses a key question about class-size reduction: “Why does it work?” Despite the many studies that show positive effects, research has yet to come up with a consistent, integrated explanation for the gains attributable to reduced class size. Several accounts have been forwarded that, by and large, have not been supported by empirical data.

One explanation, given often by researchers and practitioners, is that teachers change their teaching strategies when class sizes are reduced, providing more individualized instruction and higher quality instruction generally. This idea is intuitively appealing, but few studies show that instructional practices change spontaneously in a smaller class. Although teachers often report that small classes help them to individualize instruction, classroom observations do not support that perception. Evertson and Folger (1989) and Shapson, Eason, Wright, and Fitzgerald (1980) documented this apparent contradiction with both interview and observational data. Other studies have simply failed to find differences between teachers of small and large classes in terms of the overall structure of lessons, teaching practices, or content coverage (e.g., Bohrnstedt, Stecher, & Wiley, 2000; Cahen, Filby, McCutcheon, & Kyle, 1983; Molnar et al., 2000; Stasz & Stecher, 2002; Wang & Stull, 2000). Stasz and Stecher (2002) concluded: “This study’s findings are also consistent with research that suggests teaching practice is resistant to change and that teachers adapt their practices slowly and marginally as new materials and techniques are introduced” (p. 29).

The few changes that have been found in teaching practices are more a matter of quantity than quality. Teachers of small classes spend more time on instruction and less on classroom management or matters of discipline (Achilles, Kiser-Kling, Aust, & Owen, 1995; Bourke, 1986; Molnar, Smith, & Zahorik, 1999). These changes occur *de facto* when teachers have fewer students for whom they are responsible and when students are better behaved. However, these changes in teachers’ behavior do not seem sufficient to explain the pervasive, long-lasting academic benefits of small classes or the disproportionate benefits to low-income students.

This article summarizes theory and data on another hypothesis—that the key to the academic benefits of small classes resides in student behavior. It is proposed that students become more engaged academically and more engaged socially when class sizes are reduced, and this increased engagement in the classroom is a compelling explanation for increased learning in all subject areas. In his overview of

Project STAR, Mosteller (1995) was one of the first to speculate that student behavior explained small-class effects:

Some teachers . . . think of themselves as dealing with a start-up phenomenon. . . . Many [children] need training in paying attention, carrying out tasks, and interacting with others in a working situation. . . . [W]hen children start school, they need to learn to cooperate with others, to learn to learn, and generally to get oriented to being students. (p. 125)

This article has three sections. The first section presents a brief overview of research on two forms of student engagement in the classroom and their importance to academic achievement. The second section summarizes research on class size and student engagement in three parts: (a) research on class size and students' learning behavior, (b) research on class size and students' social behavior, and (c) research on class size and teachers' interactive styles that can affect student behavior. The third section discusses theory of group behavior that shows why small classes, like other small groups, are likely to increase student engagement. Two perspectives are presented: the effects of increasing the visibility of the individual and the effects of increasing students' sense of belonging to the group.

Student Engagement in the Classroom

Research on student engagement in the classroom has focused on behaviors such as time on task or attending to the teacher; inattentive or off-task disruptive behavior has been viewed as the other end of the continuum (disengagement). To understand the full impact of classroom organization on student behavior, however, it is necessary to distinguish engagement in learning activities from social forms of engagement. We use the phrase "academic engagement" to refer to student behaviors related directly to the learning process, for example, time on task, attentiveness, participation in learning activities, and effort and initiative taking in the classroom. "Social engagement," consisting of both prosocial and antisocial behavior, describes the nature of students' interactions with the teacher or with fellow students. Prosocial behaviors include following rules and interacting positively with the teacher as well as collaborating with other children and supporting their contributions to the class. Antisocial behaviors include withdrawing from interactions with the teacher or other students and engaging in disruptive acts that disturb the work of other students or necessitate teacher reprimand and control (Spivak & Cianci, 1987).

Both forms of engagement have consistent, strong correlations with academic performance. Studies conducted during the 1960s and 1970s demonstrated that paying attention and responding to teachers' directions were positively related to student achievement; these have been reviewed by Finn, Pannozzo, and Voelkl (1995). Several studies even demonstrated a predictive relationship between attentiveness in the early grades and later school performance. Later research continued to find strong relationships between various forms of academic engagement and students' learning (e.g., McDermott & Beitman, 1984; Peterson, Swing, Stark, & Waas, 1984). In a study of over 1,000 Grade 4 students, Finn, Pannozzo, and Voelkl (1995) found correlations of .40 to .59 between teachers' ratings of students' effort and initiative taking in the classroom and their year-end achievement scores. The correlations with ratings of inattentive behavior ranged from $-.34$ to $-.52$.

Marks (2000) studied academic engagement among elementary, middle school, and high school students. Although the purposes of the study did not include correlating engagement with academic achievement, other noteworthy findings emerged. First, engagement in instructional activities decreased as students got older. Second, the lion's share of variability in engagement (84% to 92%) was attributable to differences among students rather than to differences among classrooms or schools. Third, despite the large percentage of within-classroom variability, classroom-level influences affected student engagement. Subject matter, authentic instructional work, and social support for learning were all related significantly to students' levels of academic engagement.

Studies of students' pro- and antisocial behavior in the classroom have also produced consistent correlations with academic performance. Disruptive behavior in the elementary grades is associated with depressed academic performance (e.g., Haskins, Walden, & Ramey, 1983; Spivak & Cianci, 1987; Swift & Spivak, 1968). In the Finn, Pannozzo, and Voelkl study (1995), correlations between teachers' ratings of disruptive behavior and a battery of achievement tests were all statistically significant, ranging from $-.18$ to $-.29$. A study of junior high students (Veldman & Worsham, 1983) supported the findings from elementary grades. Observers rated the behavior of 909 students on a 24-item scale that yielded four distinct syndromes of behavior: "good," "outgoing," "rebellious," and "withdrawn." Academic performance in mathematics and reading differed significantly among the four groups. Good students and outgoing students were clearly distinct from rebellious and withdrawn students on achievement pretests, posttests, and adjusted posttests.²

Academic engagement and social engagement are associated with student race/ethnicity and socioeconomic status (SES). For example, minority students have been found to participate less fully in learning-related activities in class (Finn, Folger, & Cox, 1991; Finn, Pannozzo, & Voelkl, 1995; Lamborn, Brown, Mounts, & Steinberg, 1992) and to have higher rates of absenteeism from class and school (Bryk & Thum, 1989; Caldas, 1993) than their nonminority peers. One study showed that higher levels of engagement distinguished minority, low-SES students who were academically successful from the larger group that did not succeed in school (Finn & Rock, 1997).

Academic engagement and social engagement are strongly related to academic performance. These are precisely the processes that contribute to learning in the classroom. When a student is not engaged in learning, she or he is less likely to acquire the material presented. When antisocial behavior disrupts the teacher or other students, learning is impeded for the entire class. This article summarizes research and theory showing that one aspect of classroom organization (class size) can significantly affect student engagement and, in turn, academic achievement.

The framework for this investigation is shown in Figure 1. The focus—the relationship between class size and student engagement—is represented by the dark arrow; the effect on academic achievement is indicated by the arrow at upper right. A link between class size and teaching strategies is not solidly supported by research to date and is not included in the figure. However, teachers' morale and enjoyment of teaching are increased by small classes; this may improve their interpersonal interactions with individual students and the class (the gray arrow), further promoting academic and social engagement.

The figure also lists contextual factors that can increase or decrease the effects of reducing class size: appropriate space for learning activities with adequate tem-

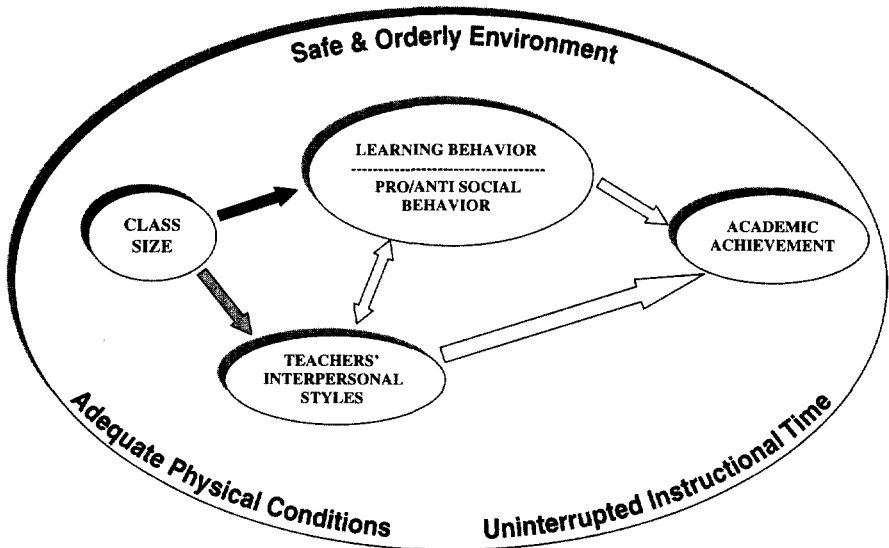


FIGURE 1. Framework used in the present investigation.

perature, lighting, and air quality; an orderly environment in which students feel safe from physical harm; and time for instruction unencumbered by significant interruptions (including pull-out programs, school reorganization, or mid-year teacher reassignments). These factors, if absent, can impede teaching and learning in any classroom. They can also prevent optimal outcomes from being realized when class sizes are reduced.

Class Size and Student Engagement: The Research

In this review, we distinguish research on three dimensions shown in Figure 1: students' learning behavior, students' social behavior, and teachers' interactive styles. Each dimension is discussed in a separate section. We identified 19 studies with sufficient information to be included in the review (Appendix A), of which 15 focused on students' learning behavior or social behavior; some studies spanned two or three of the dimensions reviewed. Several studies were conducted as part of major class-size reduction (CSR) initiatives; for background, these are described in Appendix B.

The 19 studies involved a range of methodologies, from case studies of a small number of classes to large-scale surveys. Most of the reports of quantitative research included tests of significance, but several did not. Some involved strength-of-effect measures, including percentages, and others did not. We included information on significance tests and strength-of-effect measures whenever it was available.

Class Size and Students' Learning Behavior

Of the 15 studies of class size and student behavior, 11 focused on students' learning behavior. An early study (Shapson, Wright, Eason, & Fitzgerald, 1980) investigated the effects of four class sizes (16, 23, 30, and 37) on teachers' expectations,

students' and teachers' attitudes, student achievement, teacher-student interaction, student participation, and method of instruction in the Toronto Metropolitan Schools. A total of 62 fourth- and fifth-grade classes in 11 schools participated in the 2-year study. In the first year, teachers and students were randomly assigned to one of the four class sizes in Grade 4. In the second year, teachers were again randomly assigned with the stipulation that those who had taught class sizes of 30 or 37 in the previous year had to be assigned to class sizes of either 16 or 23; similarly, students who were assigned to a class size of 16 or 37 during the first year could not be assigned to the same class size.

Attitude and opinion surveys were administered to students and teachers, and standardized achievement tests were administered to students. In addition, the Toronto Classroom Observation Schedule, composed of three components (as cited in Wright *et al.*, 1977), was used to collect information. The time-sample component provided a record of teacher behavior, student participation, and student aggressive behavior; a checklist was used to record method of instruction, subject, use of educational aids, and physical conditions; and a 45-item Likert scale was used to rate the classroom environment. Trained observers made eight half-day visits to each classroom during each year of the study.

The questionnaires assessed teachers' opinions of their classes' behavior during each year of the experiment. Responses were compared for teachers who switched from teaching a small class during the first year to a large class during the second year, and for those who switched from large to small. Teachers who changed from large classes to small classes were more likely to report that their pupils "contributed more" to class activities, $\chi^2(2) > 90, p < .001$, and "paid more attention" in class, $\chi^2(2) > 90, p < .001$.

The observations of student participation in learning activities yielded four variables: frequency of observations during which the pupil observed was participating, frequency of observations during which the pupil was participating nonverbally, frequency of observations during which the pupil was participating verbally, and frequency of observations during which the pupil had no task to perform. Differences in class size on these variables were tested using one-way analyses of variance (ANOVAs). Results of the analyses showed no significant differences on any of the student participation variables as a result of class size. That the authors did not find any significant results may be attributable to design features. The observations were made in Grades 4 and 5 rather than in earlier grades, and students had only attended small classes for 1 year and then changed to larger classes. Patterns of engagement behavior may be relatively stable by this time and difficult to change. Further, only one class size (16) fell in the category "small" as practiced in recent interventions.

Cahen, Filby, McCutcheon, and Kyle (1983) investigated the impact of reducing class size midway through a school year in four Grade 2 classes in two schools, one in rural Virginia and one in an inner-city district in California. At the beginning of the school year, enrollments were 20 in the two classes in Virginia and 35 in the classes in California. Midway through the school year, one third of the students from each class were randomly assigned to a newly created class such that the two original classes and the newly created class included students with a range of achievement levels. Observers in Virginia visited classrooms three or more times each week over the course of the school year; observers in California

visited classrooms a minimum of once a week. Detailed field notes were recorded on tape by each observer for subsequent translation. Also, observers in Virginia only conducted observations during three reading and two mathematics lessons prior to classes being reduced and four reading and four mathematics lessons once classes had been reduced (Cahen et al., 1983, p. 186). The observers used a time sampling method to code the engagement behavior of six target students in reading and six in mathematics as "engaged-academic" (student is paying attention to an academic task), "engaged-procedural" (student is paying attention to some activity that supports the academic task [i.e., passing out papers]), "engaged-rules" (student is paying attention to activity related to rules or classroom management), "waiting for help," "off-task" (student is not paying attention but is being quiet and not disruptive), "misbehaving" (student is not paying attention and doing something potentially disruptive), and "down time" (student is waiting because he or she has nothing to do). Finally, teachers kept journals throughout the year, and researchers conducted semiformal interviews with the teachers and principals.

The "engaged-academic," "engaged-procedural," "waiting for help," "off-task," and "down time" variables each describe some aspect of students' learning behavior. Median rates for each student were calculated and pooled across subject matter and teacher. Results of *t* tests comparing the small-class phase and large-class phase revealed a significant difference in the average rates students were engaged in academic tasks in favor of small classes ($t = 4.1, p < .001$). During the small-class phase, average engagement in academic tasks was 18.0% higher than during the large-class phase (large classes: $M = 56.6\%$, $SD = 13.4\%$; small classes: $M = 74.6\%$, $SD = 13.4\%$). For down time, the difference between the small- and large-class phases was also significant in favor of small classes ($t = -3.71, p < .002$). The average "down-time" rate decreased from 13.9% in the large-class phase ($SD = 13.0\%$) to 0.0% in the small-class phase ($SD = 0.0\%$). Large but non-significant differences were found between the rates of time spent waiting for help and off-task in the small- and large-class phases. Results for the engaged-procedural variable were not reported owing to the low reliability of this measure.

Project Prime Time (1984–1986) was designed to reduce class sizes in primary classrooms (K–3) throughout Indiana to 18 students (or 24 with a teacher aide). Classes with more than 18 students with one teacher, or more than 24 with an aide, were designated as large. In their 1986 report, Chase, Mueller, and Walden discussed the effects of Prime Time on student achievement and the impact on teachers' instructional practices and student behavior. Data were collected in the form of teacher surveys, teacher and principal interviews, parent surveys and interviews, and classroom observations. A total of 349 Grade 1 and Grade 2 teachers completed the teacher survey. A portion of the survey asked teachers to compare their classes with classes prior to Prime Time on a number of Likert-type items with possible responses of *a great deal more*, *somewhat more*, *about the same*, and *somewhat less*. For the item "pupils are on-task," 77.0% of the teachers responded "a great deal more" or "somewhat more." Researchers also compared the responses of teachers of large classes with those of teachers of small classes on the same items using chi-square tests. For "pupils are on-task," significantly more small-class teachers than large-class teachers answered "a great deal more" or "somewhat more" (84% and 73%, respectively; $\chi^2 = 9.06, p = .01$).

Several studies of learning behavior were conducted in conjunction with Project STAR. In Year 3 of STAR, when students were in second grade, Everston and Folger (1989) investigated class size and teaching practices. Trained observers observed a total of 52 Grade 2 classrooms in 13 schools during reading and mathematics lessons. Observers recorded teacher-to-student and student-to-teacher contacts in behavioral, academic, or procedural contexts and took descriptive notes to gather information. Observations were conducted once in the fall and once in the winter of 1987 during one reading and one mathematics lesson. Among the variables developed from the coding sheet were the following: student-initiated questions, student-initiated comments, student-initiated academic contacts, students definitely on-task, students off-task, and students waiting. Average percentages for small and regular classes were computed separately for reading and mathematics for each variable and then compared through t tests (although t values were not reported). For the variables percentage of student-initiated questions and percentage of student-initiated comments, there were no statistically significant differences between small and regular classes in reading or mathematics. However, students in small classes initiated a higher percentage of academic contacts in mathematics than did students in regular classes, although the percentages were both small (7.3% vs. 4.1%; $p = .03$; $d = .58$). No significant difference was found in reading. In addition, a significant difference in percentage of students definitely on-task was found favoring small classes in reading (87.4% vs. 83.2%; $p = .04$; $d = .40$) but not in mathematics. No significant differences between small and regular classes were found in reading or mathematics for percentage of students off-task or percentage of students waiting.

These results partially support the premise that class size affects student behavior. The authors concluded that, in small classes, each student is likely to get a turn more often during lessons, and students initiate more contacts with teachers. Unfortunately, the results of this investigation are weakened by the authors' incomplete explanation of how the variables were created.

Finn, Fulton, Zaharias, and Nye (1989) investigated the carryover effects of small classes in Project STAR using Grade 4 data from students who had been in the same class type for Grades 1–3. Grade 4 teachers were asked to complete the Student Participation Questionnaire (SPQ; Finn, Folger, & Cox, 1991) for up to 10 students currently in their class. The SPQ consists of 25 items rated according to frequency of occurrence, from *never* to *always*. The items are combined into three behavioral subscales, two of which relate to learning behavior, Minimally Adequate Effort (e.g., “pays attention in class”) and Initiative Taking (e.g., “does more than just the assigned work”), and one that relates to social behavior, Nonparticipatory Behavior (e.g., “is withdrawn or uncommunicative”). In total, teachers in 58 of the 76 STAR schools completed questionnaires for 2,207 students. Differences in teacher ratings for students who had previously attended small, regular, and regular-aide classes were analyzed using multivariate analyses of variance (MANOVAs). Statistically significant differences ($p < .0001$) were found among the three class types on the set of three participation measures. Univariate tests of the small-regular class contrast revealed statistically significant results in favor of students who had attended small classes on the effort and initiative subscales ($p < .05$, $d = .12$ and $p < .01$, $d = .14$, respectively). Students who had attended small classes in Grades 1–3 displayed significantly more effort and initiative, as rated by their Grade 4 teachers.

When STAR students reached Grade 8, 1,481 students were rated on a shortened 14-item form of the SPQ. Each student was rated by his or her mathematics and English teachers, and the responses were averaged. Minimally Adequate Effort, Initiative Taking, and Nonparticipatory Behavior scale scores were formed. Again, MANOVA was used to compare students who had attended small, regular, and teacher-aide classes. No significant differences were found among the means (Voelkl, 1995). The behavior differences found in Grade 4 were not apparent in Grade 8, even though academic achievement was affected through Grade 8 (Finn, Gerber, Achilles, & Boyd-Zaharias, 2001).³ The short-term impact of small classes was diminished after students spent several years in large classes. The dynamics of larger classes in Grade 4 and beyond may have counteracted the earlier behavior gains.

A reduced-class initiative prompted by STAR, Success Starts Small (SSS), was targeted to two elementary schools in High Point, North Carolina, from October 1993 to June 1994. The goal was to compare early elementary grades (K–2) in two Title I–eligible schools. One school used a traditional Title I pull-out model, while the other school used Title I resources to create small classes. Average class sizes were 23 in the traditional school and 14 in the school that created small classes. Information was gathered via interviews, structured classroom observations, informal classroom visits, pretest and posttest scores on versions of the California Achievement Test (CAT), and discipline referrals.

The structured classroom observations involved coding interactions using the PIT method developed by French and Galloway (n.d.). Interactions between teachers and students were classified every 4–5 seconds into one of three categories: "personal" (not related to academic activities or school), "institutional" (related to daily classroom routines), or "task" (related to academic activities). Interactions were also coded as having an "individual" focus, a "group" focus, or a "mixed" focus (if there was more than one focus). Observations were conducted in fall 1993 (pre) and May 1994 (post). In an analysis of the Grade 1 observations, Achilles, Kiser-Kling, Aust, and Owen (1995) found that a consistently high percentage of interactions in the school implementing small classes were task related—approximately 82% (pre) and 84% (post) of all interactions. In contrast, the percentages of task-related interactions in the school implementing traditional classes were 79% (pre) and 67% (post). The percentages of task-related interactions focusing on individual students rather than groups were also different between the two schools. In the school implementing small classes, the pre and post percentages were 51% and 44%, respectively. The corresponding percentages in the school implementing traditional classes were 31% and 33%.

Unfortunately, no tests of significance were conducted to compare pre-to-post observations or differences between the schools. Despite the absence of statistical tests, the authors concluded that a greater percentage of interactions took place between teachers and individual students in small classes than in large classes. Further, the interactions of teachers and students in small classes were more often related to academic or learning activities than were the interactions of teachers and students in large classes.

Egelson and Harman (2000) reported the impact of small classes on instructional strategies in Grade 1–3 classrooms in Burke County, North Carolina. Five schools were chosen for observation based on "end-of-grade" tests in reading and mathematics so that they represented the continuum of test results in the county.

Class sizes had been reduced to 18 or fewer students as part of a district-wide initiative begun in 1991–1992. Observers spent 15 minutes, unannounced, in 8 to 10 small classes in each school and recorded classroom interactions using the School Observation Measure (Smith, Ross, Alberg, & Lowther, 1999). Observers also scored the level of student attention as “low,” “medium,” or “high,” based on an estimate of the percentage of time spent by students in learning activities. It is unclear whether observers rated each individual student, a select number of students, or generalized across the class as a whole. Ratings from the classes in each school were then averaged to form an overall rating of student attention for the school. In 4 of the 5 schools, the average rating of student attention levels was “high”; in the fifth school, it was “medium.” Unfortunately, interpretation of these findings is hindered by the absence of a comparison with nonreduced classes.

In 1997, Maier, Molnar, Percy, Smith, and Zahorik reported the results of the first year of Wisconsin’s SAGE (Student Achievement Guarantee in Education) program. As part of the evaluation, data were collected on classroom organization, instructional practices, and student behavior through interviews with principals and teachers, teacher questionnaires, teacher activity logs, a shortened 16-item version of the SPQ (used by Finn, Fulton, Zaharias, & Nye, 1989, in Project STAR), and classroom observations. Kindergarten and Grade 1 teachers completed a participation questionnaire for each student in their class in fall 1996 and spring 1997.

Using principal components analysis for the SPQ items, Maier et al. (1997) identified two factors: On-Task Behavior and Active Learning Behavior. The On-Task Behavior scale included items such as “pays attention in class” and “acts restless, is often unable to sit still.” The Active Learning Behavior scale included items such as “participates actively in discussions” and “asks questions to get more information.” The authors documented an increase in average ratings from fall to spring on both scales. For the On-Task Behavior scale, teachers’ average ratings of students increased approximately 8%, from 70.50 in the fall ($SD = 17.48$) to 77.22 in the spring ($SD = 16.59$). For the Active Learning Behavior scale, teachers’ average ratings increased 6%, from 69.85 in the fall ($SD = 17.84$) to 78.39 in the spring ($SD = 16.84$).

The SAGE evaluation has contributed to our understanding of the academic effects of small classes. However, the behavior data suffer from the absence of a comparison with larger classes. Also, SAGE utilized alternate classroom configurations under the rubric “small classes,” including large team-taught classes. As a result, it is difficult to attribute these results to the small-class setting.

The California class-size initiative included an assessment of student learning behavior. Information was gathered through a 49-item teacher questionnaire and through intensive case studies of classroom processes in third-grade classes that used teacher interviews, observations, teacher logs, videotaped lessons, and samples of student work (CSR Research Consortium, 2000). Questionnaire results were reported for 531 third-grade teachers who taught in reduced-size classes and 105 who taught in nonreduced classes. Items on the questionnaire included statements about students’ behavior. Teachers rated both the general behavior of students during the most recent school day and specific behaviors during the most recent language arts lesson. Teachers chose from among five options ranging from “no students” to “all students.” The evaluators converted the responses to percentages for comparison.

Analyses revealed differences in favor of small classes, but none of these differences were statistically significant. In comparison with teachers in nonreduced classes, teachers in reduced classes reported a lower percentage of students as "off task (but not disruptive) for more than 5 minutes" during language arts lessons (28.2% vs. 24.9%). Teachers in reduced classes reported a higher percentage of students who "completed the whole lesson for the day" during language arts lessons than did teachers in nonreduced classes, but the difference was small (72.9% vs. 69.4%). A nonsignificant difference was also reported between reduced and nonreduced classes in the percentage of students who "sought assistance from another student regarding schoolwork" during the most recent school day.

New research in England has examined the relationship of class size to teaching behavior, student attention/off-task behavior in class, and peer relations (Blatchford, in press). The classes considered were 39 "reception" classes for students 4–5 years of age. The observational study involved 235 children in 21 small and 18 large classes drawn from a larger 3-year longitudinal study. A large group of schools was identified that had both small (fewer than 20 students) and large (more than 30 students) reception classes. Target schools were randomly selected, and educators were asked whether they were willing to participate. Within each school, the names of six (plus two alternate) children from each class were selected at random for observation. The observers were trained in the use of an observation schedule that included the following categories: "setting" (individual work, group work, play or free time, and transition times), "subject" (mathematics, language, other task work, topic, and other activities), and "social modes"⁴ (including teacher-child contact, child-teacher contact, child-child contact, and "not interacting"). Each social mode was divided into subcategories (see Table 1).

Each child was rated in 10-second intervals over a 5-minute period during academic activities (i.e., language, mathematics, crafts, painting, free play). Ultimately, the goal was for each of the six children to be observed five times over the course of a school day on 3 separate days. However, as a result of student absences and scheduling variability, each child was observed for an average of 69 minutes, or roughly 14 five-minute intervals.

Differences between small and large classes on each subcategory were analyzed by ANOVA.⁵ The seven subcategories related to learning in the "child-teacher" category revealed statistically significant results ($p < .05$ to $p < .001$) in favor of small classes. In comparison with children in large classes, children in small classes were more frequently observed initiating interactions with the teacher, as well as responding and attending to the teacher during interactions. Children in small classes frequently engaged in continuing interactions with the teacher, as evidenced by the significant results for the "continued" variable. In terms of the content of the interactions, children in small classes were more frequently observed to be interacting about academic tasks or procedures than were their counterparts in large classes. Finally, children in small classes were observed to be off-task less frequently than were children in large classes.

Similarly, separate analyses of the "child-child" subcategories revealed statistically significant results in favor of small classes for one of the two variables related to learning ("task"; $p < .01$). The means for small classes on this variable were *lower* than the means for large classes. This may indicate that students in

TABLE 1

Social modes, subcategories, and p values for comparison between large and small classes^a

Social mode	Subcategories	<i>p</i>
Teacher–child contact	One-to-one setting	<.001
	Group setting	<.001
	Whole-class setting	<i>ns</i>
	Focus of teacher’s attention ^(T)	<.001
	Teacher teaching content ^(T)	<.001
	Teacher preparing a task ^(T)	<.001
	Teacher being passive ^(T)	<.001
	Teacher interacting about classroom procedure ^(T)	<.001
	Teacher interacting socially ^(T)	<.01
	Teacher’s behavior unclear ^(X)	<i>ns</i>
Child–teacher contact	Child responding to teacher ^(L)	<.05
	Child initiating contact ^(L)	<.05
	Child attending to teacher ^(L)	<.001
	Child engaged in continued interaction ^(L)	<.001
	Child’s behavior unclear ^(X)	<i>ns</i>
	Child interacting about a task ^(L)	<.001
	Child interacting about classroom procedure ^(L)	<.01
	Child interacting socially ^(S)	<.05
	Child interacting inappropriately ^(S)	<i>ns</i>
	Child off-task ^(L)	.05
Child’s purpose unclear ^(X)	<.01	
Child–child contact	Child interacting about a task ^(L)	<.01
	Child interacting about a procedure ^(L)	<i>ns</i>
	Child interacting socially ^(S)	.001
	Child “mucking about” (fooling around) ^(S)	<.001
	Child being aggressive ^(S)	<i>ns</i>
	Child helping ^(S)	<.01
Child’s behavior unclear ^(X)	<.001	
Not interacting	Child working alone on task ^(L)	<.01
	Child working alone on procedure ^(L)	.01
	Child actively off-task ^(S)	.001
	Child passively off-task (daydreaming) ^(L)	<.001
	Child in audience ^(X)	<.05
	Child’s behavior unclear	<.001

Note. (T) indicates teacher behavior to be discussed in a separate section; (L) indicates student learning behavior to be discussed in this section; (S) indicates student social behavior to be discussed in a separate section; (X) indicates not discussed in this article.

^aAdapted from Tables 2–5 of Blatchford (in press), with permission of the publisher.

large classes rely more on each other for assistance because the teacher is not as accessible. There was no statistically significant difference for the “procedure” subcategory.

Analysis of the “not interacting” subcategories revealed statistically significant results in favor of small classes on all three variables related to learning. The “task

involved," "procedure," and "off-task (passive)" subcategories were significant ($p < .01$ to $p < .001$). Students in small classes spent less time working on tasks without interacting with the teacher and their peers; they were more likely to be interacting and collaborating with the teacher or their peers during tasks. Conversely, students in small classes spent more time alone engaged in completing classroom procedures; they had greater independence in day-to-day and nonacademic routines. Finally, students in small classes were less likely to be passively off-task or inattentive than were students in large classes.

In sum, students in small rather than large classes were observed to initiate and engage in contacts with teachers more frequently, to engage in interactions with teachers directly related to learning more frequently, and to be off-task less frequently. Students in small classes spent less time working alone on tasks without interacting with the teacher or their peers and spent more time alone completing classroom procedures and daily routines. These results support our hypothesis that small classes affect students' learning by changing the nature of interactions and relationships in the classroom among students and between students and the teacher. While some differences were not statistically significant, none of the 12 variables significantly favored large classes.

Conclusions. The research investigating the impact of class size on learning behavior is extremely diverse. The studies reported here used a wide range of definitions of student learning behavior and several methods to investigate the role of class size. Student learning behavior was operationalized as on-task academic behavior (i.e., "pays attention in class," "engaged in academic activity," or "off-task but not disruptive") or behaviors thought to be related to learning (i.e., "effort," "initiative," and "active learning"). These behaviors were assessed using single indicators as well as more complex measures for which several indicators were assessed and aggregated into one index. Six of the 11 studies relied on classroom observations. In five of the six, observers rated the learning behavior of individual students within each class; in the sixth, observers developed an overall or average rating of the learning behavior of the class. The remaining five relied on teacher reports of learning behavior; in three, teachers rated individual students in their class, and in two, teachers rated the overall behavior of students in the class. Adding to the diversity are the range of class sizes termed "small" and the length of time students attended small classes (from one-half year to 4 years).

The mixture of approaches made the task of finding consistencies difficult. Nonetheless, the 11 studies showed a positive impact of smaller classes on students' learning behavior regardless of the methods or measures employed. The two strongest studies (Blatchford, in press; Finn, Fulton, Zaharias, & Nye, 1989) reported results that were statistically significant on all but one variable, with effect sizes ranging from small to moderate. These studies had clear definitions of learning behavior, used well-constructed measures, and employed rigorous methodologies. Between them, they covered a continuum of students' learning behaviors, including effort and initiative taking in the classroom and actual time on task.

Of the remaining studies, one reported significant results in favor of small classes (Chase, Mueller, & Walden, 1986). One study reported results that were nonsignificant but favored small classes (Voelkl, 1995). Four others reported mixed significant and nonsignificant results but always favoring small classes

(Cahen, Filby, McCutcheon, & Kyle, 1983; CSR Research Consortium, 2000; Evertson & Folger, 1989; Shapson et al., 1980). Three studies did not conduct tests of significance (Achilles, Kiser-Kling, Aust, & Owen, 1995; Egelson & Harman, 2000; Maier et al., 1997), but results consistently favored small classes. Altogether, results favored small classes on at least 38 of the 42 measures employed.⁶ While some individual changes were small, no measure in any study was statistically significant in favor of larger classes.

There is also variability in the results according to grade level. Eight of the 11 studies focused on students in Grades K–3. Results favored small classes in all eight studies. Of the studies that examined behavior beyond Grade 3, little if any impact of class size was found. Shapson, Wright, Eason, and Fitzgerald (1980) examined student behavior in Grades 4 and 5; students in this study attended small classes for 1 year at most. Finn et al. (1989) found behavior differences in Grade 4 for students who attended small classes in K–3, but Voelkl (1995) found that the behavior effects did not persist through Grade 8. Grade 4 effects were not surprising since STAR participants attended small classes all day long for up to 4 years. Even then, the large-class environment of Grades 4–8 seemed to thwart the continuation of positive learning behaviors.

Several factors may account for the weaker findings in later grades. For example, early small classes socialize students to the work of schooling (Mosteller, 1995); behavior may be harder to affect as students grow older. Also, the norms of larger classes encountered in later grades may eclipse the impact of smaller classes, and being engaged in learning may take different forms in middle school years and beyond, especially when students move from class to class for different subjects and spend more learning time outside of class. In general, the behaviors assessed in elementary grades may be less pertinent to learning in the later grades.

The lack of rigor of a number of the studies reviewed limits the conclusions that can be drawn. In addition to the paucity of information about the definition and measurement of student learning behavior, some reports of observations failed to explain how variables were operationalized or how observers were trained. The duration of the observations was sometimes unclear and, in some studies, too short to be reliable; for example, Egelson and Harman (2000) reported making 15-minute observations. No study took into account the length of time students had attended small classes as a result of students transferring in and out of schools, an essential feature of the intervention being studied. Finally, 3 of the 11 studies failed to conduct simple tests of significance that may have strengthened the conclusions, and 2 lacked any comparison group. Of the studies that reported statistically significant results, only 3 included a strength-of-effect measure. Several others actually provided the information needed for effect sizes but did not compute them. This may be especially important because several of the differences reported appeared to be small.⁷ Despite all of this, the trends were highly consistent.

Further study of the impact of class size and classroom organization generally on students' behavior is needed. Related questions have not been addressed at all, for example, how the effect of class size on behavior is moderated by classroom contextual factors such as the adequacy of the physical environment, students' sense of safety in the classroom, or the orderliness of the class (see Figure 1). These questions should be addressed through research that employs well-defined and reliable measures of learning behavior, that makes clear whether individual student

behavior or class behavior is being studied, and that also gives attention to features of the class (e.g., composition of the student body, grade level).

Class Size and Students' Social Behavior

Research on the effects of class size on students' social behavior has focused largely on antisocial behaviors such as disruptiveness and misbehavior. Fewer studies have focused on changes in prosocial behaviors, for example, cooperating/collaborating with other children or following rules. We found a total of 10 studies that reported on class size as it related to students' social behavior. All 10 reported the impact on antisocial behavior, and 6 also reported the impact on prosocial behavior.

Dennis (1986) investigated the impact of small classes in Grade 2 on a number of variables, including student achievement, attendance, self-concept, and behavior, in the Nashville (Tennessee) schools. Using "experimental," control, and blind control groups, the author compared students who were enrolled in small classes (15 or fewer students) with students who were enrolled in regular classes (approximately 25 students) and a sample of students from across the district. Students in the "experimental" group ($n = 85$) were enrolled in one of six small second-grade classes from one school; they were not randomly assigned to small classes as in Project STAR. Students in the control group ($n = 88$) were enrolled in one of four second-grade classes from a school with similar racial, socioeconomic, and geographic composition. Students in the blind control group ($n = 85$) were selected from the second-grade population of 43 schools in the Nashville district. Students in the experimental and control groups were matched according to gender, race, SES, birth date, and reading scores.

Students' antisocial behavior was assessed indirectly, through corporal punishment records maintained by the principal of each school and the central data processing department of the district. Corporal punishment consisted of paddling, administered mostly by male principals to male students and occasionally by female principals to female students. Data were collected on the number of corporal punishment incidents from the first 6 months of the 1985–1986 school year for each group. Principals indicated that corporal punishment incidents were most often the result of students' "disruptive behavior" or "refusing to obey." Statistical tests revealed significant differences between the experimental and control groups in favor of the experimental (small class) group ($z = -3.35, p < .01$). Significant results were also found for the experimental and blind control (comparison) group comparison in favor of small classes ($z = -1.67, p < .05$). No effect sizes were reported. The author concluded that the lower rate of corporal punishment incidents in small classes indicated a positive impact on student behavior. However, the author also collected data on the corporal punishment records for students in the experimental and control groups for the 2 years prior to the small class intervention. In reviewing these data, it appears that these two groups may have differed significantly in their antisocial behavior, with students in the small-class group having fewer incidents from the outset.

Chase, Mueller, and Walden (1986) reviewed data collected through teacher surveys in Project Prime Time. A total of 349 Grade 1 and Grade 2 teachers completed the teacher survey about their teaching experience and education, instruc-

tional practices, and perceptions of their classes. A portion of the survey asked teachers to compare Prime Time classes with their prior classes. One item was an indicator of students' prosocial behavior ("the classroom is peaceful and quiet"), while a second item indicated antisocial behavior ("there are discipline problems"). A significantly higher percentage of teachers of small classes reported their classrooms to be "peaceful and quiet" after Prime Time than did teachers of large classes (72% and 56%, respectively; $\chi^2 = 13.01, p < .01$). The chi-square test for discipline problems was not statistically significant. Since small classes in Prime Time used a variety of organizational models, however, neither of these results can be attributed to small classes alone.

Three studies conducted in conjunction with Project STAR included some investigation of students' social behaviors. Everston and Folger (1989) investigated the impact of class size on antisocial behavior of Grade 2 students during reading and mathematics lessons in 52 classrooms. Observers coded teacher-to-student and student-to-teacher contacts in terms of "percentage of behavioral contacts," "amount of disruptive behavior," and "amount of inappropriate behavior." A significant difference was found in favor of small classes for percentage of behavioral contacts in mathematics (7.3% vs. 11.4%; $p = .04; d = .45$) but not in reading. For amount of disruptive behavior and amount of inappropriate behavior, a 5-point scale was used (1 = lowest frequency, 5 = highest frequency). Comparisons between small and large classes were made across reading and mathematics lessons. Small but significant differences were found in favor of small classes for both amount of disruptive behavior (1.6% vs. 2.0%; $p = .01; d = .36$) and amount of inappropriate behavior (1.6% vs. 2.1%; $p < .01; d = .61$). That is, across both lessons, there were fewer contacts between teachers and students that focused on student behavior in small classes, and students in small classes displayed somewhat less disruptive and inappropriate behavior than did students in full-size classes.

In an investigation of the carryover effects of small classes, Finn, Fulton, Zaharias, and Nye (1989) asked teachers to rate the behavior of Grade 4 students who had been in small classes or full-size classes from Grade 1 through Grade 3. As mentioned earlier, each teacher rated a random sample of 10 students in his or her class using the SPQ, which includes an "antisocial" nonparticipatory behavior scale composed of four items (e.g., "needs to be reprimanded" and "annoys or interferes with peers' work"). A statistical test of the difference between students who had been in small classes and students who had been in full-size classes revealed significant results in favor of students in small classes ($p < .05, d = -.11$). In this study, the behavioral benefits of small classes were found to carry over to Grade 4, after all students returned to full-size classes.

Johnston (1990) investigated teachers' perceptions of teaching and their students' pro- and antisocial behavior in STAR small classes. During May of each year (1986–1989), exit interviews were conducted with teachers involved in the study that year. More than 1,000 K–3 teachers were interviewed over the 4-year period. The interviews with the kindergarten teachers were used to identify and define common themes that were used to categorize the interview protocols. Fourteen categories emerged related to teachers' perceptions of teaching, including one indicating prosocial behavior and one indicating antisocial behavior. The "social climate" category indicated positive interactions between students, teachers and

students, and students' willingness to cooperate. The "classroom management" category identified perceptions related to antisocial behavior, including students' problem behaviors.

In terms of social climate, teachers consistently perceived students in small classes to be more cohesive as a group and generally more cooperative, supportive, tolerant, and caring, in contrast to their experiences with regular-size classes. Teachers expressed that children in small classes seemed more willing to take risks themselves, and also "encouraged classmates to try, to risk, and would not accept less than a good effort from their peers" (p. 13). Changes teachers noted in the social climate of the classroom in smaller classes are illustrated by the following:

Second grade and third grade small-class teachers observed that the children appeared to be less inhibited, less afraid of being wrong and that they volunteered to answer questions more often than students in full-size classes. One teacher observed, "They feel safe with their ideas and they're not going to be put down." Teachers described children in small classes as more curious, enthusiastic, and eager to participate than were children in their regular size classes. Several teachers noted this was particularly the case in their low achieving reading group. (p. 11)

However, some Grade 2 and Grade 3 teachers noted that when students remained with the same classmates for 4 years, they were not always receptive to new classmates joining them.

Teachers expressed that classroom management was much easier with a small class and there were generally fewer behavior problems. However, they did not identify specific problem behaviors that occurred less frequently. Teachers in small classes seemed to perceive management to be easier because having fewer children made it "easier to be aware of potential problems before they became problems" (Johnston, 1990, p. 14).

Since the completion of Project STAR, other studies have examined the effects of reduced class size on social behaviors. Achilles, Kiser-Kling, Aust, and Owen (1995) studied the impact of Success Starts Small (SSS), a project involving two elementary schools in High Point, North Carolina. The researchers reviewed discipline referrals made by Grade 1 teachers in the school that implemented small classes. Discipline referrals were made by teachers for students being disrespectful, refusing to do work, being disruptive, hitting/fighting, making threats, stealing, and having weapons, all indicators of antisocial behavior.

Discipline referrals decreased consistently in the 2 years after small classes were implemented. There was a 26% drop from the first year with full-size classes to the second year and a 50% drop from the second to the third year. In this study, it was not possible to determine whether the decrease was the result of fewer discipline problems or whether teachers were finding ways to address discipline issues directly in the classroom. Either outcome is beneficial, however. Unfortunately, no discipline data were collected in the school using full-size classrooms, so the authors were not able to compare the two schools.

Grade 1 teachers participating in SSS were interviewed by Kiser-Kling (1995) about the immediate benefits of a class size of 15 students. Their comments echoed Johnston's (1990) finding of an increased sense of community in small classes and explained in part why disciplinary problems declined. The teachers reported "more

time to give 'troubled kids' the attention they so desperately need, reducing the likelihood of their becoming a discipline problem [*sic*]" (p. 150). When the same teachers later taught large classes (about 24 students), they reported "more student conflicts in the classroom" as a common theme.

The evaluation of Wisconsin's SAGE program (Molnar, Smith, & Zahorik, 1999) included case studies of teaching in three separate schools. One Grade 1, Grade 2, and Grade 3 class was observed in each school. Each school used a different SAGE classroom configuration: "regular" (a small class with 15 students), "shared space" (two small classes in a classroom that has been temporarily divided), or "team" (30 students with 2 full-time teachers). The case studies involved four formal observations of each class during reading and mathematics lessons, additional informal observations, and three formal interviews of the classroom teachers.

The nine teachers reported consistently that discipline was easier and student misbehavior was less of an issue than in large classes. In addition, they identified changes in the overall classroom atmosphere. The three teachers in the "small" classrooms expressed that students seemed to be "voluntarily compliant" by appearing more focused on tasks and being more willing and eager to help their peers. All three teachers felt they knew and understood their students in the SAGE classrooms better than students they had previously in larger or traditional classrooms. In turn, this created a more "personal, relaxed, family-like atmosphere" (Molnar, Smith, & Zahorik, 1999, p. 100).

In the evaluation of California's class-size reduction initiative (CSR Research Consortium, 2000), the responses of 636 third-grade teachers to the 49-item questionnaire were examined. Teachers were asked to rate the overall behavior of students in their class during the most recent school day, as well as specific behaviors of students during the most recent language arts lesson. Teachers chose one of five options that ranged from "no students" to "all students." Responses were converted to percentages by the evaluators. Four items related to students' social behavior (three antisocial items, one prosocial item) were included.

Differences were found between teachers' ratings of reduced and nonreduced classes in the percentage of students "competing with another student for the teacher's attention in class" during the most recent school day, 29% of students in reduced classes compared to 42% of students in nonreduced classes ($p < .01$). Significant differences were also found in the percentage of students "engaged in exclusionary behavior (i.e., forming cliques)" during the most recent school day, 19% of students in reduced classes compared to 26% of students in nonreduced classes ($p < .01$).

The same group of teachers also rated the behaviors of students during the most recent language arts lesson. A small but significant difference was found between reduced and nonreduced classes on the percentage of students who "disrupted the work of other students," 19% in reduced classes compared to 23% in nonreduced classes ($p < .01$). No significant difference was found between reduced and nonreduced classes on the percentage of students who "helped another student complete his/her work."

Finally, teachers who had taught both reduced-size and nonreduced classes responded to an open-ended question about the most important differences

between the two. Teachers provided 1,902 separate responses that were coded into 24 distinct categories. Of the top 10 categories listed in the evaluation report, the first three pertained to the ability to individualize attention and instruction to meet students' needs. However, "easier class discipline" emerged as the fourth most important difference, with 20% of all teachers listing this in their responses. "Easier to manage class" and "fewer problem students, students better behaved" were also among the top 10, with 14% and 10% of teachers including these themes in their responses, respectively.

Two recent evaluations examined students' social behavior in some depth, one using classroom observations and the other using focused teacher reports. Blatchford (in press) investigated the impact of class size on interactions among students and between students and teachers in 39 British reception classes (students 4–5 years of age). Observers coded the interactions into four major categories, each with multiple subcategories (see Table 1). Three of the major categories included subcategories related to social behavior.

Differences were found between small and large classes in terms of students' interactions with teachers, students' interactions with their peers, and students' behavior when they were not interacting with others. In terms of students' interactions with teachers, students in small classes engaged in a greater number of social interactions with teachers (e.g., interactions not related to academics or school routines) than students in large classes ($p < .05$). No significant difference was found between small and large classes in the number of inappropriate interactions between students and teachers. That is, students in small and large classes engaged in roughly the same number of interactions with their teachers that were not related to instruction.

Differences were also found in students' interactions with other students. Students in small classes engaged in fewer social interactions than did their peers in large classes ($p = .001$). Students in small classes were also found to help other students less often than did students in large classes ($p < .001$). "Helping other students" was conceptualized in this study as nonlearning behavior (e.g., "helping a peer tie shoelaces"). Thus, the author interpreted this result as favoring small classes by suggesting that students in large classes spend more time on their own and less with the teacher; we view this conclusion as ambiguous. Students in small classes also engaged in less "mucking about" (fooling around) than did students in large classes ($p < .001$), but no significant difference in aggressive behavior was found between students in small and large classes. When students were not interacting with the teacher or their peers, students in small classes were less likely to be actively off-task (disruptive) than were students in large classes ($p = .001$).

In general, however, students in small classes engaged in more social interactions with their teachers but did not behave inappropriately. Similarly, students in small classes engaged in fewer social interactions with their peers and "fooled around" less than did students in large classes. Students in small classes were also less likely to be disruptive. These results paint a picture of social behavior in small classes that is generally more positive than in larger classes.

An in-depth evaluation of the CSR initiative in Buffalo, New York, included questionnaires completed by teachers asking them about their experiences and their perceptions of the program (Finn, Forden, Verdinelli, & Pannozzo, 2001). Teach-

ers in both reduced and nonreduced classes in the same grades in the same schools responded to the questionnaires. In all, 213 teachers completed questionnaires.

Questionnaire items asked about the “manageability” of the class, students’ classroom behavior, and the sense of community in the classroom. The manageability scale consisted of seven items such as “When students behave improperly, it is difficult to maintain control of the classroom” and “I am able to develop a close, friendly relationship with every student in my class.” The classroom behavior scale, composed of six items, included “Students follow classroom rules” and “Some students are disruptive.” The sense of community scale, with four items, included “Students encourage each other to do well” and “There is substantial conflict among students.” The manageability and behavior scales related more directly to students’ antisocial behavior, while the sense of community scale reflected prosocial behavior. Tests of significance were performed on the mean score differences between small and large classes for each of the three scales. In each case, a statistically significant difference was found in favor of small classes: manageability, $t(210) = 4.59, p < .0001, d = .75$; behavior, $t(210) = 3.03, p < .003, d = .50$; and sense of community, $t(210) = 3.23, p < .002, d = .53$. On average, teachers in small classes perceived classroom management to be easier, classroom behavior to be more positive, and there to be a greater sense of community than did teachers in large classes.

Conclusions. Like the research on students’ learning behavior, this body of research has relied on a wide range of definitions of student social behavior and a mixture of methods to investigate the impact of class size. Student social behavior was more frequently conceptualized as negative or antisocial behavior than as positive or prosocial behavior. Antisocial behavior was defined as either disruptive to the teacher or students (e.g., “needs to be reprimanded,” “annoys or interferes with peers’ work”) or in terms of discipline issues (e.g., corporal punishment records or referrals made to the principal). Prosocial behavior was defined as relating to the community or classroom atmosphere, operationalized in ratings such as “the classroom is quiet and peaceful,” “students engaged in exclusionary behavior,” and “students assisted other students.” All 10 studies reviewed included measures of antisocial behavior; 6 also included a measure of prosocial behavior.

Studies of students’ social behavior relied heavily on teacher reports; 6 of the 10 studies relied on teacher reports, and 2 studies utilized observer ratings of students. The remaining 2 studies used records kept by the schools.

Overall, the 10 studies showed a positive impact of smaller classes on students’ social behavior—decreasing antisocial behavior and promoting prosocial behavior. Three studies reported results that were completely statistically significant (Dennis, 1986; Finn, Forden, Verdinelli, & PannoZZo, 2001; Finn, Fulton, Zaharias, & Nye, 1989), and one reported results that were statistically significant on all but one variable (Blatchford, in press). Three studies reported mixed significant and nonsignificant results but always favoring small classes (Chase, Mueller, & Walden, 1986; CSR Research Consortium, 2000; Everston & Folger, 1989). Of the remaining studies, one did not conduct tests of significance, but the means also favored small classes (Achilles, Kiser-Kling, Aust, & Owen, 1995); two studies were qualitative (Johnston, 1990; Molnar, Smith, & Zahorik, 1999). Altogether, of

24 measures of pro- and antisocial behavior, 17 were statistically significant in favor of small classes, although some effects were small. One was significant but ambiguous as to direction, and 5 were nonsignificant; no test of significance was conducted on the remaining measure. With one possible exception, no measure in any study was statistically significant in favor of larger classes.

The pattern of results for antisocial behavior was highly consistent. Despite a range of definitions of antisocial behavior, these studies demonstrated clearly that reducing class size reduces antisocial behavior. Students in small classes were less likely than students in large classes to fool around, engage in inappropriate behavior, or disrupt the work of other students, and they had fewer referrals for discipline problems. Since most studies were based on teacher ratings, it is not clear whether the decrease in disruptive behavior is the result of a change in students' behavior, teachers redefining what they considered disruptive, or both. The observational studies suggest that it was a combination of both. It is clear by all accounts, however, that antisocial behavior decreases when class size is reduced.

The evidence of a positive impact of class size on students' prosocial behavior is less complete. Smaller classes appear to promote an atmosphere in which students are more supportive and caring about each other. However, very little research has focused on prosocial behavior. Rather, researchers have included more positive behaviors as a contrast to the focus on antisocial behavior. It is important that further research focus on a range of students' prosocial behaviors and classroom atmosphere. In addition to well-designed quantitative studies, qualitative studies can help address questions such as the following: (a) How do students exhibit positive relationships with one another in small classes? (b) How do students at risk respond to the more positive behavioral environment of small classes? and (c) If students feel a greater sense of belonging in small classes, how is that exhibited?

Class Size and Teachers' Interpersonal Styles

The conceptual model for this review proposes that smaller classes also affect teachers' morale and enjoyment of teaching, which in turn affect students' levels of engagement. This can occur in several ways: (a) if teachers' demeanor is better and students perceive that the teaching/learning situation can be enjoyable; (b) if teachers are less punitive, allowing students to exhibit a greater breadth of normal behavior; and (c) if teachers get to know each pupil better, increasing students' sense of community in the classroom. Research on this issue is scanty. We located only nine studies that approach this topic, each from a different perspective. Although the results are consistent, more work is needed in this area.

In an observational study in Australia, Bourke (1986) investigated the relationships among class size, teaching practices, and student achievement in 63 Year 5 (elementary) classes in 33 schools. Class size was determined by dividing the number of students in the room by the number of teachers in the room for each 5 minutes of time observed; values were then averaged. Sizes ranged from 12 to 33 students ($M = 25.2$). Observers recorded various teaching practices over a 12-week period, including the amount of "nonacademic management," "noise tolerated in the class," and direct interactions between teachers and students (academic and nonacademic combined).

Significant negative correlations were found between class size and “teachers directly interacting with students” and between class size and “amount of noise tolerated in the class” ($r = -.27$ and $r = .34$, respectively); p values were not reported. As class size increased, the amount of time teachers were observed to spend in direct interaction with students decreased, and teachers tolerated more noise in their classrooms. In addition, a significant positive correlation was found between nonacademic management and class size. As class size increased, the amount of time teachers were observed to spend on nonacademic management increased. Unfortunately, the report did not define “nonacademic management,” which could include activities such as correcting papers, setting up lessons, and so on, as well as managing behavior.

Two studies in conjunction with Project STAR included some investigation of teachers’ interpersonal styles; one is Johnston’s (1990) investigation of teachers’ perceptions of teaching and the classroom environment. In interviews with approximately 1,000 K–3 teachers, one salient category to arise was “interpersonal relations.” Teachers in small and regular-aide classes noted improvements in their interactions with students, as well as changes in their knowledge of children, their families, and their home backgrounds.

Kindergarten through third grade teachers reported that more time was now available to *listen* to children, to get to know their personal lives and concerns. Conversely, teachers also perceived that children knew more about the teacher as an individual with a history, interests, and a life outside of school. Teachers reported feeling more like a part of the class. (p. 12)

Small-class teachers in particular commented that “children were more willing to approach the teacher . . . they frequently initiated conversation with teachers about personal matters” (p. 13). Other changes were also noted. Teachers reported “lower levels of noise in the classroom and their own increased tolerance for noise and movement” (p. 14). Finally, teachers “felt more proactive and less reactive” (p. 14) in their approach to managing student behavior in the classroom.

In a study of STAR teachers in Grades 2 and 3, Boyd-Zaharias and Pate-Bain (2000) investigated differences in teaching practices between high-quality and lower quality teachers. Teachers were classified as high quality if their classes’ scores on the Stanford Achievement Test fell in the top 15% of the score distribution; they were compared with other teachers whose classes scored in the bottom 50% of the distribution. Sixty-five teachers in each group were interviewed and observed teaching. Observers rated them as poor, fair, good, or excellent on a number of criteria adapted from the Concepts of Effective Teaching (Northwest Regional Educational Laboratory, 1984). For example, “standards for classroom behavior are explicit” and “incentives and rewards for students are used to promote excellence” relate to teachers’ style of classroom management, while “personal interactions between teacher and students are positive” relates to teachers’ interpersonal styles. Information about how observers arrived at their ratings was not provided by the authors.

Although the study was not explicitly a study of class size, the authors noted that a high proportion of high-scoring classes were small (over half of the top 10% of kindergarten classes and 78% of the top 10% of classes in Grade 3). Thus, the comparison of high-quality and low-quality teachers was essentially a comparison between small and larger classes. Highly significant differences were found

between the percentages of teachers rated as excellent in the high and low scoring groups. A significantly higher percentage of teachers in the high achievement group were rated as excellent on "standards for behavior are explicit," as compared with teachers in the low achievement group (86% vs. 56%; $p < .001$). Similarly, a higher percentage of teachers in the high scoring group than in the low scoring group were rated as excellent on "personal interactions between teacher and student are positive" (88% vs. 56%; $p < .001$). Finally, a sizable but nonsignificant difference was found between the percentages of high- and low-quality teachers rated as excellent on "incentives and rewards for students are used to promote excellence" (82% vs. 63%).

In North Carolina's *Success Starts Small* (Achilles, Kiser-Kling, Aust, & Owen, 1995), a group of Grade 1 teachers had the unusual experience of teaching small classes (about 15 students) for several years and then being assigned to larger classes (about 24 students). Kiser-Kling (1995) interviewed these teachers about the differences they perceived. Teachers reported "more personal teacher-student conversations" and "more time to get to know each student and their families" as immediate benefits of small classes and "I don't know the children as well" as the most common outcome of the larger classes.

The evaluation of Project SAGE (Molnar, Smith, & Zahorik, 1999) included qualitative analyses of teacher practices and the classroom environment. As mentioned earlier, case studies were conducted of teaching in nine classes using three different SAGE classroom configurations: regular (a small class with 15 students), shared space (two small classes in a classroom that has been temporarily divided), and team (30 students with 2 full-time teachers). The case studies involved four formal observations of each class during reading and mathematics lessons, additional informal observations, and three formal interviews with the teachers.

Results of the teacher interviews and observations showed that teachers spent more time teaching because discipline was not an issue. According to the authors:

Discipline problems disappear in reduced size classes not only because inappropriate behavior is instantly recognized in a small class and can be given a response with no delay and because the teacher-student proximity reduces its occurrence, but because in a reduced class size inappropriate behavior is re-defined. As several teachers revealed, in a small class students are given more freedom. Many behaviors not tolerated in a large class, because of the problems they create, such as walking around the room, are acceptable in a small class. Further, there is also less misbehavior because students' greater understanding in small classes causes them to be less confused and, subsequently, better behaved. (pp. 100-101)

Stasz and Stecher (2002) evaluated the effects of class-size reduction in California on teachers' teaching practices, use of time, and management strategies through case studies of 16 third-grade teachers during the second and third years of CSR. Case studies included interviews with principals and teachers, content surveys, classroom observations of mathematics and language arts lessons (two times per year, one of which was videotaped), daily activity logs, and collections of students' work (p. 5). The authors reported results for case studies of two teachers as they made the transition from teaching in large classes during the first year to teaching in small classes during the second year.

Prior to CSR, the authors described the classroom management practices of the first teacher as actively managing noise and attention levels. "For example, she said 'shh' repeatedly and almost continuously throughout the lesson . . . and jingled a little bell when the noise level grew too loud. Her actions seldom became oppressive, but were constantly present" (p. 17). However, the transition to teaching in the small class had a definite impact on her classroom management practices. In the reduced class, she did not quiet students all the time. "The noise in her reduced-size class was apparently more tolerable and thus she did not feel the need to continually remind students to be quieter" (p. 27).

The second teacher's overall management style did not change substantially. Her style was described as very directed and controlling in both settings, leading to few disruptions in either setting. Disruptions were dealt with in a blunt and direct manner. For example, "Sam, did you want to spend more time with me? I don't think so" (p. 23). In the small-class setting, "she was still clearly in control of the action in her classroom, but she also seemed to provide more positive feedback to students in her reduced-size class" (p. 27). For both teachers, "the amount of time and energy devoted to discipline, order, and transitions declined with the small classes" (p. 27).

Krieger (2001) investigated the impact of class size on teacher and student interactions in Grades K–2 in a suburban Louisiana school district. The study involved four small classes (18 students or fewer) and seven large classes (24 or more) in three elementary schools. Principals nominated teachers who were perceived to be effective. Each classroom was videotaped on three occasions during language arts and mathematics lessons for a minimum of 1.5 hours during the spring of 1999. The videotapes were analyzed in a two-stage process to determine emergent categories of verbal interaction and were scored for a priori categories of nonverbal interaction.

Seven categories of verbal interactions emerged, out of which two were related to teachers' interpersonal styles: positive attention (e.g., "I like the way everyone is working") and negative attention (e.g., "Why would you ever get out of a chair without pushing it under?"). On average, 5.9% of all verbal interactions of teachers in small classes provided positive attention, while only 3.6% of verbal interactions in large classes provided positive attention. And, on average, 10.2% of verbal interactions in small classes provided negative attention, in contrast to 17.8% for teachers in large classes. Nonverbal behaviors of teachers were seen to affect communication in both small and large classes. In general, teachers of small classes were found to use more facial expressions and more eye contact than teachers of large classes; teachers of large classes were found to use posture and proximity more than teachers of small classes. Unfortunately, it is difficult to understand the full implications of this study because of the absence of clear definitions of the behavior categories or a clear summary of differences between small and large classes.

The recent evaluation by Blatchford (in press) represents the most focused investigation of teachers' interpersonal styles as related to class size. Observers coded interactions between students and teachers in small and large reception classes (4- to 5-year-olds). One set of observations, termed "teacher-child contact," included three subcategories related to teachers' interpersonal behavior: "target child is focus of teacher's attention," "teacher being passive," and "teacher interacting socially."

Differences were found in favor of small classes on each of the three subcategories. In comparison with teachers in large classes, teachers in small classes engaged in more interactions with students where the target child was the focus of the interaction as opposed to in the audience ($p < .001$). Teachers in small classes also engaged in more interaction with students where they took a passive role (e.g., listening to a student read, observing a student work) than did teachers in large classes ($p < .001$). Finally, teachers in small classes engaged in a greater number of social interactions (i.e., talking about life outside the classroom) than did teachers in large classes ($p < .01$). These results paint a picture of teachers in small classes who spend more time engaged in direct interactions with each student and interacting socially with each student. Teachers of small classes also spent more time in a passive role with students, interacting without the need to guide students academically or behaviorally.

Finally, an in-depth evaluation of the CSR initiative in Buffalo, New York, included interviews with 38 teachers in the program (Finn, Forden, Verdinelli, & Pannozzo, 2001). The majority of teachers interviewed taught in small classes, but a small number of teachers taught in inclusion classes. Teachers were asked to describe the most positive features or strengths of the program and make recommendations they felt would enhance the program in subsequent years. They identified "teaching students in small groups and having more contact with each student" and "fewer discipline problems and improved student behavior" most often as benefits of small classes. The third most salient theme identified as a strength of the program was an improved sense of community in their classrooms. Nine teachers reported a variety of social benefits in terms of student-student and student-teacher relationships. In addition, "teachers were able to know their students on a more personal basis" (p. 47). Those teachers who could compare their experience with a small class to prior experience with larger classes reported that they knew more about students' lives outside of school.

Conclusions. The nine studies reviewed here begin to demonstrate relationships between class size and teachers' interactive styles. Of the three mechanisms postulated, the studies confirm two: that teachers get to know each student more intimately in small classes and that their tolerance for a broader range of student behaviors is increased. The results of the research are consistent, and we conclude tentatively that these effects do occur. None of the studies reviewed addressed the third possibility: that teachers' enjoyment of teaching is increased in small classes and students' enjoyment of learning increases as a result.

Unfortunately, the methodologies of the studies limit the generalizations that can be drawn. Too many studies used small numbers of classes and teachers' comments that may be socially biased as sources of evidence. In addition, the studies provide little help in explaining these processes in depth. No study focused broadly on teachers' interactive styles and how they were affected by classroom organization. Moreover, no study examined the possible impact of class size on teachers' interactive style and, in turn, on student behavior. This area, as much as any, is in need of further research. It is especially important given that research has not found that teachers make substantial changes in their instructional practices when class sizes are reduced.

Toward a Theory of Class Size and Engagement

The studies reviewed show that smaller classes can affect students' social and academic behavior. To understand how this occurs, we examined the work of psychologists and sociologists on group size and member participation. Two principles, termed "visibility of the individual" and "sense of belonging," were drawn from social science theory and research; we view them as components of a theory linking class size to student behavior.

Although little of this work has been applied to young children, the principles are very similar to descriptions of small classes given by teachers and observers (for examples, see Finn, Forden, Verdinelli, & Pannozzo, 2001, and "Teachers Speak Out" in Finn & Wang, 2002). Both principles present plausible hypotheses for explaining student behavior and should be tested in classroom research.

Visibility of the Individual

In the classroom, the visibility principle asserts that students in a small class experience increased pressure to participate. They are "on the firing line" by being more visible to the teacher and may be called upon at any time to answer questions or to participate in a class activity; the pressure to be prepared and to respond appropriately is greater than it would be in a large class. Students in a small class cannot easily avoid being noticed, and the teacher cannot readily ignore any pupil(s) even if she or he would like to. It is no surprise to find students highly engaged in learning activities under these conditions.

Psychological theory has used the concept of visibility and its antithesis, anonymity, to explain the reduced participation of members of larger groups in general. The effects on motivation and on active participation have been studied in terms of "diffusion of responsibility" and "social loafing," respectively. Each of these may explain why small classes increase student participation.

Diffusion of responsibility. Diffusion of responsibility refers to motivational levels: "When individuals are part of a group, they tend to feel less responsible" (Forsyth, 1999, p. 460). The ideas behind this construct evolved after an event in New York City in 1964 in which a woman was murdered in front of 38 witnesses; none summoned help. Since that time, dozens of studies have shown that individuals feel their responsibility to help those in distress is greater if they are the only person present than if a group is present. Latane and Nida (1981) reviewed 10 years of research on group size and helping behavior. Of thousands of individuals studied, about three fourths helped people in distress when alone, but only about one half helped when they were part of a group.

This is not to say that a classroom is a crisis situation (at least not usually), or that students are bystanders. But the basic premise behind diffusion of responsibility—as group size increases, the felt responsibility to act decreases (see Baron & Byrne, 1997)—may apply in many situations. To date, it has not been demonstrated in classrooms or among young children; further research is needed to determine whether the principle applies to students in classes of different size.

Social loafing. The motivational factors suggested by diffusion of responsibility have not been tested broadly. However, the actual behavior of individuals in groups

has been studied extensively under the rubric "social loafing." Social loafing is the "reduction in effort by individuals working in groups" (Forsyth, 1999, p. 288). It is common to so many situations, tasks, and populations defined by gender, age, and cultures that it is truly a pervasive characteristic of groups. The extent of social loafing among members of a group is related directly to group size.

The concept of social loafing originated at the beginning of the 20th century with respect to collective tasks—tasks in which individuals work together to produce a single result. Typical outcomes that were studied included rope-pulling ("tug-of-war") and applauding. Researchers observed that groups become less efficient—because each participant exerts less effort—as group size increases (Latane, Williams, & Harkins, 1979; Petty, Harkins, Williams, & Latane, 1977; Ringlemann, 1913). Each study also indicated a "group-size plateau," or a point above which adding members to a group did not further decrease individual effort. The plateau was different for each task and set of work conditions. Subsequent experiments confirmed that social loafing occurs with a wide variety of tasks and when tasks are performed coactively, that is, when "individuals work in the real or imagined presence of others, but the inputs of different people are not combined" (Karau & Williams, 1995, p. 135). This describes what happens in the classroom.

Several theories have been forwarded to explain why social loafing occurs, including social impact theory, based on the premise that "increasing the number of persons in . . . [a] target group should diminish the pressures on each individual [to participate] because the impact is divided among the group members" (Latane, Williams, & Harkins, 1979, p. 830; see also Karau & Williams, 1993). Two factors work in tandem to affect social loafing: anonymity and evaluation apprehension. "People performing simple tasks . . . often expend less effort when their contributions are hidden from others' evaluative scrutiny" (Forsyth, 1999, p. 289). Individual effort is likely to be reduced (a) when an individual is less noticeable, (b) when she or he believes that a contribution will have little impact on group functioning, and/or (c) when she or he believes that the contribution is less likely to be evaluated or rewarded. These all occur to a greater extent in larger groups than in smaller groups.

When individual contributions are clearly identifiable—that is, visible—social loafing is reduced or eliminated (see Hardy & Latane, 1986; Kerr & Bruun, 1981; Sanna, 1992). Visibility increases automatically when the size of the group is decreased. Psychologists are not unanimous about whether evaluation apprehension results from visibility or whether it is a separate phenomenon. But studies have shown that (a) the presence of others reduces the fear of having one's individual contributions evaluated, and (b) social loafing is increased when individuals believe their contributions are not comparable or cannot be evaluated by a common standard (see, for example, Gibb, 1951; Harkins, 1987; Harkins & Jackson, 1985; Szymanski & Harkins, 1993).

These conditions pertain directly to the classroom, although classroom research on social loafing is limited. It has been demonstrated that larger classes permit students to reduce their visibility ("hide in the crowd") by sitting in less conspicuous places. Adams and Biddle (1970) identified a T-shaped "action zone" of seats in larger classes that accounts for much of the interaction between students and teachers: "The 'fringe dwellers' on the outskirts of the room are not directly involved in

the educational transaction” (p. 51). Kashti, Arieli, and Harel (1984) observed interactions in three classrooms with 29, 25, and 16 students. In the two larger classes, groups of students sat where they would not be particularly noticeable and avoided calling attention to themselves. By being less visible, they reduced the likelihood that their individual contributions would receive any attention or be evaluated negatively. No such groups were formed in the smallest class. Small classes make it difficult for students to become anonymous and are likely to reduce the extent of social loafing.

Sense of Belonging

From a “sense of belonging” perspective, smaller groups may encourage member participation (a) because smaller groups are more unified in their purposes and actions than are larger groups and (b) because individual members often feel that they are more closely affiliated with the group, receiving guidance and support from other group members. These principles were described at the beginning of the 20th century by sociologist Georg Simmel (see translation by Wolff, 1950); they have been studied in contemporary times in terms of “group cohesiveness” and “psychological sense of community,” respectively. In the classroom setting, smaller classes are likely to be more unified in their purpose, and, as observed frequently, teachers are likely to have closer relationships with each individual student and students tend to be more supportive of each other.

Group cohesiveness. Group cohesiveness has been defined as “the value of the rewards that participation in the group provides for its members” (Homans, 1974, p. 118), that is, the extent to which members are attracted to the group (not just to individuals who are part of it). Cohesive groups tend to have high morale, and members tend to support and coordinate efforts with one another (Forsyth, 1999; Shaw, 1976). Two questions are central to applying the concept to classrooms: (a) Does cohesiveness increase the likelihood that members will participate in group activities? and (b) Do smaller groups tend to have greater cohesiveness than larger groups?

That group norms affect the behavior of individuals is a well-established principle of human behavior. It is also confirmed that the press to conform is greater in cohesive groups. Experiments dating from the 1940s (e.g., French, 1941) through the past decade (e.g., Hogg, 1992) have demonstrated this principle with activities ranging from forming opinions to interpreting ambiguous stimuli and completing questionnaires. Shaw (1976) summarized the findings of years of research on the topic as follows: “Groups characterized by friendliness, cooperation, interpersonal attraction, and similar indications of group cohesiveness exert strong influences upon members to behave in accordance with group expectations” (p. 201).

Are smaller groups more cohesive than larger groups? Simmel (as translated by Wolff, 1950) wrote: “As the size of the group increases, the common features that fuse its members into a social unit become ever fewer. . . . The variety of persons, interests, events becomes too large to be regulated by the center” (pp. 397–398). Recent research is consistent with this proposition. Mullen and Copper (1994) summarized studies of group size and cohesiveness and concluded that as group size increases, members’ liking for the group and performance are both reduced. Their own meta-analysis found this to hold for real groups as well as for “artifi-

cial" groups (in which subjects assume that they are part of a group). The authors concluded that the impact of group cohesiveness "is generally stronger among smaller groups" (p. 220).

It has also been found that as group size increases, the proportion of *nonparticipants* increases. Bales et al. (1951) found that as the number of participants in a group discussion increased from three to eight, the number of participants who took part declined, and others ceased to participate altogether. Bray, Kerr, and Atkin (1978) observed the same phenomenon and coined the term "functional group" to describe those members who continued to participate and carried the lion's share of the whole-group effort. Further, large groups may splinter and promote the formation of subgroups (Shaw, 1976, p. 156) whose attitudes and activities may run counter to those of the primary group. This may be the phenomenon observed by Kashti, Arieli, and Harel (1984): In classes of 29 and 25, informal groups of students formed with distinct patterns of classroom participation or nonparticipation; in the smaller class of 16 students, "no definable subgroups were formed" (p. 175).

Group cohesiveness occurs in classrooms to the extent that students encourage each other to engage in learning activities—by showing that this is "the thing to do" and by giving support and encouragement to their classmates' efforts. Interactions such as these can be rare or nonexistent when children first begin school but can increase over time *under favorable conditions*. Even in the early grades, there is anecdotal evidence that students in small classes are more supportive of one another and interact more positively than do children in large classes (Achilles, 1999; Finn, Forden, Verdinelli, & Panno, 2001). In larger classes, students may actually discourage their classmates' efforts and/or form splinter groups that interfere with teaching and learning. The cohesiveness hypothesis needs to be examined further in relation to class size and student behavior.

Psychological sense of community. It is widely held by psychologists that people have a pervasive need to belong to social groups (Baumeister & Leary, 1995). Sarason (1974) coined the phrase "psychological sense of community" (PSOC) to represent "the perception of similarity to others, an acknowledged interdependence with others, a willingness to maintain this interdependence by giving to or doing for others what one expects from them, the feeling that one is part of a larger dependable and stable structure" (p. 157). Researchers who study PSOC have concluded that community membership serves four major purposes for the individual: "membership," the feeling of belonging; "influence," the sense of mattering to the group; "integration and fulfillment of needs," the feeling that the community will meet the needs of the individual; and "shared emotional connection," an emotional bond with other members of the community (McMillan & Chavis, 1986). Some research on PSOC has been conducted in school settings.

Peer groups generally, and classes in school in particular, are important affiliation groups for many students, serving all four of these purposes and, in turn, affecting classroom behavior. Research has shown that students' sense of belongingness in school (Goodenow, 1993; Wehlage et al., 1989) and identification with school (Voelkl, 1995) are associated with a range of outcomes, including motivation to perform well, engagement in learning activities, pro- and antisocial behav-

ior in the classroom, and academic performance (Bateman et al., 2000; Osterman, 2000; Royal & Rossi, 1996).

A key question is whether smaller classes are more likely than larger classes to satisfy students' needs for a sense of community. Sociological research concludes that "ingroup bonds tend to be stronger the smaller group size is . . . as are strength of affective ties and satisfaction" (Blau, 1977, p. 135). No research has been conducted to answer this question in school settings, but Bateman (2002) argues that this is the case. In her own research, Bateman et al. (1998) found that students in "Schools for Thought" (SFT) classrooms had a greater sense of community than did students in traditional classrooms. Students in SFT classes had high levels of participation in learning activities and a great deal of interaction with classmates. Even among traditional classrooms, the class with the highest degree of PSOC was one in which students were encouraged to be highly participatory. These processes occur more commonly in small classes, suggesting that class size may be a determinant of PSOC.

Bateman's argument focuses on the four dimensions identified by McMillan and Chavis (1986), incorporating the findings of previous class-size research. Membership is enhanced in small classes because students are "more active and frequent participants in the learning process" (p. 10). Influence is increased when students develop strong feelings of academic and social competence, in turn enabling them to have positive effects on their peers. In terms of integration and need fulfillment, Bateman posited that "students' needs for autonomy, for engaging in challenging activities, and for social support and the presence of a social comfort zone can be better met through the structure and activities facilitated by a small classroom" (2002, p. 12). This occurs because teachers have more time to attend to the needs of individual students and are able to give more latitude to a range of student learning styles and classroom behaviors. Finally, small classes are likely to facilitate shared emotional connections. Research⁸ has shown that students in small classes exhibit lower levels of antisocial behavior. The opportunity to support classmates and to collaborate with them is increased.

Class size may also affect teachers' sense of community with the class. According to principles of group dynamics, leaders in larger groups are likely to be more active participants and other members tend to be less active, and as groups get larger, leaders tend to communicate more with the group as a whole and direct fewer comments to individual members (Forsyth, 1999; Shaw, 1976). Indeed, teachers of small classes report that they get to know all of their students better in a small-class setting (e.g., Achilles, 1999; Finn, Forden, Verdinelli, & Pannozzo, 2001). If they interact more often and more directly with individual students, an increased sense of community among students and teachers as a group may result.

As with cohesiveness, little or no research has directly examined the development of PSOC in elementary-grade students. The age or stage of life at which PSOC develops and becomes an influential motivating condition has not been documented, nor has the age at which a class group can satisfy an individual's need for group membership. Further work is needed to understand the role of students' PSOC in shaping classroom behavior, in small classes and in classes of all sizes.

Summary and Implications

This investigation was an attempt to explain why students in the elementary grades benefit academically by attending small classes. Achilles and Finn (2002)

identified four sets of factors that may be affected when class sizes are reduced: students' learning activities and behavior, teaching practices, classroom conditions, and "others," including teachers' morale and parents' involvement. In this review, we have focused on student behavior—that is, engagement—because of consistent data showing that, when class sizes are reduced, students become better behaved and more intensely involved in learning activities.

In our model (Figure 1), teachers play an important role to the extent that they affect student engagement through increased enjoyment of teaching, tolerance for a wider range of student behavior, and increased interpersonal interactions with students.⁹ The effects of class size on students and teachers may be moderated by classroom conditions such as safety, light and air quality, and number of interruptions. Unfortunately, little if any research has examined the interaction of these classroom conditions with class size or with classroom organization in general.

The studies of class size and student engagement we reviewed were extremely diverse—both small-scale and large-scale studies, both quantitative and qualitative studies, and studies of different quality. Several methodological problems were found in the research, including sometimes inadequate measurement of student behavior and an overreliance on teacher reports. Thirteen of the 19 studies reviewed made some use of teacher reports, including interviews, questionnaires regarding the class or individual students, and surveys about their impressions of their classrooms. Given the context of these studies, it is unlikely that any teachers would indicate "larger is better," introducing possible bias into their responses. Also, several studies did not conduct statistical tests of significance, and those that did often neglected to discuss the strength of effects. In some cases, this made it difficult to ascertain whether differences between small and large classes were educationally meaningful.

Nevertheless, the research consistently supports two conclusions: Students in small classes in the elementary grades are (a) more engaged in learning behaviors and (b) display less disruptive behavior than do students in large classes. The research is consistent with two other conclusions but, to date, does not provide complete evidence: (a) Students in small classes exhibit increased prosocial behavior relative to students in large classes, and (b) improvements in learning behavior in the early grades appear to fade out in later grades. The latter occurs as students spend more years in large classes. That the academic benefits continue through later grades is yet to be fully explained. Finally, the research suggests that teachers' interpersonal styles benefit from class-size reduction, and this, in turn, affects students' engagement in the classroom. Because of the paucity of research in this area, we have only a partial picture of the teacher's role.

Although rarely applied to classrooms, social science theories provide a basis for understanding why student behavior is affected by class size. We organized this theory according to two principles: "visibility of the individual" and "sense of belonging." The first principle explains how behavior changes when members of groups become more visible because they have fewer co-members. Students in a small class are constantly "on the firing line," promoting increased attention and participation; students in larger classes can be more anonymous and reduce their interactions with the teacher.

Psychologists and sociologists have documented two ways in which this occurs. First ("diffusion of responsibility"), individuals who are part of a larger group tend

to feel less responsible for completing group tasks. And second (“social loafing”), individuals who are part of a larger group put forth less effort than do individuals in smaller groups. In some settings, social loafing has been shown to “plateau”; that is, loafing increases as groups get larger up to a certain point. This may happen when classes reach a size threshold (e.g., 17 or 20 students) or when classes are truly overcrowded (e.g., 30 or more students). Social loafing has been found to decrease when contributions are likely to be evaluated individually. If teachers in smaller classes undertake more individual assessments and give more feedback to students, student effort is likely to increase as well.

The second principle—sense of belonging—explains the interplay of individuals and groups with which they identify. That group norms influence members’ behavior is well established. Social scientists have demonstrated that small groups tend to be more cohesive than large groups, thus affecting individual behavior in a more focused way. If this principle applies to the classroom, small classes are likely to exert more influence on each student to become and remain engaged. Larger groups may splinter and promote the formation of subgroups with their own identities and behavior. Thus, large classes may encourage groups of students to exhibit behavior at odds with the goals of the teacher and school, for example, inattentive and/or disruptive behavior.

The individual’s need to feel that she or he is part of the group has been termed “psychological sense of community.” PSOC promotes a sense of belonging especially when group activities are seen as important and when the individual is given support and acceptance. This is more likely to occur in a small class (Bateman, 2002). Research indicates that teachers of small classes accept a wider range of normal student behavior. Further, student-teacher relationships undoubtedly benefit because small classes allow teachers to spend more time with each student instead of seeing the class as “a sea of faces.” Of all the theory reviewed here, however, the sense-of-belonging principle has received the least attention from researchers.

High-Priority Research Needs

This article raises significant questions about classroom and school organization that are not understood adequately by researchers or practitioners. Rather than catalogue all possible research issues, we highlight four that are pressing. Well-designed observational studies are needed, in addition to other research approaches, to address each issue.

First, what factors promote student engagement, that is, positive learning behaviors and prosocial behavior? Engagement needs to be viewed as a “dependent variable,” not just another independent variable. The role of engagement in learning is clear. We need to focus more on how to engage students in class and in school generally—especially students who are withdrawn, inattentive, or disruptive. Small classes can help, but research has yet to tell us about other aspects of classroom organization, curriculum organization, and instructional practices that enhance student engagement.

Second, what is the impact of class size and other organizational features on different groups of students, particularly older students or students at risk for school failure? Despite research showing that small classes have the greatest benefits for students at risk, despite the clear connection between engagement and perfor-

mance, and despite the prevalence of dysfunctional classroom behavior among low-income, minority students, none of the 19 studies cited in this review made comparisons based on race/ethnicity, SES, or even gender. Several studies collected this information to use in achievement comparisons but did not use it when examining student behavior (e.g., Bourke, 1986; Maier et al., 1997). Because improved learning behavior and social behavior may provide a pathway to increased academic achievement, this research focus is a must.

Early summaries of class-size research (e.g., Glass & Smith, 1978; Robinson, 1990) gave some attention to class size in later grades. However, Project STAR and other recent initiatives have turned attention exclusively to the primary grades. Some justification for this is provided by the finding that the greatest academic benefits accrue to students who enter small classes early (kindergarten or first grade) and who remain in small classes for 3 or 4 years (Finn et al., 2001). However, the issue of student behavior becomes more significant in every ensuing grade. The dynamics discussed in this review and the mechanisms forwarded to explain them may be just as relevant to middle-grade or high school students, or more so. Research that parallels the work described here, but with older students, is an important priority.

Third, research is needed to understand the role of the contextual variables in our model (Figure 1) as they affect student engagement. Engagement is enhanced when the physical conditions of the classroom (e.g., air quality, noise level) are adequate, when interruptions are kept to a minimum, and when students feel safe from harm.

The interaction of contextual features with class size is a two-way street. On the one hand, class size (or other organizational features) may affect the context. For example, Prout (2000) has shown that air quality falls to below acceptable standards by mid-day in large classes but remains adequate in small classes. And the increased cohesiveness of small classes may reduce students' fear of ridicule or bullying by classmates. On the other hand, contextual factors "moderate" the impact of small classes. The effectiveness of small classes or, for that matter, any classroom intervention may be seriously compromised if instructional time, physical conditions, and feelings of safety are inadequate. Despite the obvious importance of these factors, research on their effects is lacking.

Finally, this article has introduced four mechanisms to explain why small classes affect students' academic and social engagement: diffusion of responsibility, social loafing, group cohesiveness, and psychological sense of community. Limited classroom-based research to date has examined these processes indirectly, for example, by studying similar dimensions but using different terminology. An organized conceptual framework that synthesizes these processes into a single model, giving appropriate consideration to the empirical findings, would be useful for researchers and practitioners alike.

Research directly linking these mechanisms to schools and classrooms is virtually nonexistent, and many questions remain to be answered. Among them are the following: How can the constructs best be operationalized and observed in classroom settings? At what age does theory based on adolescents and adults become germane to young children? To what extent do principles of group dynamics, studied mostly in small groups, apply to groups as large as 20 or 30? And, ultimately, to what extent do these principles explain student behavior in small classes or, indeed, in classes of any size?

Notes

Work on this article was performed with the support of a grant from the Spencer Foundation titled “A Study of Class Size and At-Risk Students.” An earlier version was presented at the annual meeting of the American Educational Research Association, New Orleans, April 2002. We are grateful to Harvey Pines for identifying reference material on individual and group behavior.

¹Neither Project STAR nor any of the studies reviewed by Glass and Smith (1978) or Robinson (1990) was included in Hanushek’s (1998) review.

²Unfortunately, neither standard deviations nor effect sizes were reported in the article.

³It may have continued further. The state testing program for all students ended in Grade 8, however.

⁴This is the term used by Blatchford, but the variables fit the definition of learning behavior as used in this review.

⁵Scores for the major categories, obtained by summing the subcategory scores, were also analyzed by ANOVA. Because they comprise both academic and social engagement, and because they are somewhat redundant after the subcategories were analyzed, we do not discuss them separately. All differences were statistically significant.

⁶Shapson et al. (1980) did not provide the means needed to make this determination for nonsignificant findings.

⁷As noted in the text.

⁸Reviewed in the first sections of this article.

⁹We acknowledge that, in the bigger picture, the role of the teacher is still more important. For example, despite the promising findings about class size, small classes cannot counteract teaching practices that fail to recognize and reward—or even alienate—students who make sincere efforts to participate.

References

- Achilles, C. M. (1999). *Let's put kids first, finally: Getting class size right*. Thousand Oaks, CA: Corwin Press.
- Achilles, C. M., & Finn, J. D. (2002). The varieties of small classes and their outcomes. In M. C. Wang & J. D. Finn (Eds.), *Taking small classes one step further* (pp. 121–146). Greenwich, CT: Information Age.
- Achilles, C. M., Kiser-Kling, K., Aust, A., & Owen, J. (1995, April). *A study of reduced class size in primary grades of a fully Chapter 1-eligible school: Success Starts Small*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Achilles, C. M., Nye, B. A., & Zaharias, J. B. (1995, April). *Policy use of research results: Tennessee's Project Challenge*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Adams, R. S., & Biddle, B. J. (1970). *Realities of teaching*. New York: Holt, Rinehart & Winston.
- Bales, R. F., Strodtbeck, F. L., Mills, T. M., & Roseborough, M. D. (1951). Channels of communication in small groups. *American Sociological Review*, 16, 461–468.
- Baron, R. A., & Byrne, D. (1997). *Social psychology* (8th ed.). Boston: Allyn & Bacon.
- Bateman, H. V. (2002). Students' sense of community: Implications for class size. In M. C. Wang & J. D. Finn (Eds.), *Taking small classes one step further* (pp. 63–75). Greenwich, CT: Information Age.

- Bateman, H. V., Bransford, J. D., Goldman, S. R., & Newbrough, J. R. (2000, April). *Sense of community in the classroom: Relationship to students' academic goals*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Bateman, H. V., Goldman, S. R., Newbrough, J. R., & Bransford, J. D. (1998). Students' sense of community in constructivist/collaborative learning environments. In *Proceedings of the Twentieth Annual Meeting of the Cognitive Science Society* (pp. 126–131). Mahwah, NJ: Erlbaum.
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497–529.
- Blatchford, P. (in press). A systematic observational study of teachers' and pupils' behaviour in large and small classes. *Learning and Instruction*.
- Blau, P. M. (1977). *Inequality and heterogeneity*. New York: Free Press.
- Bohrnstedt, G. W., Stecher, B. M., & Wiley, E. W. (2000). The California class size reduction evaluation: Lessons learned. In M. C. Wang & J. D. Finn (Eds.), *How small classes help teachers do their best* (pp. 201–225). Philadelphia: Temple University Center for Research in Human Development and Education.
- Bourke, S. (1986). How smaller is better: Some relationships between class size, teaching practices, and student achievement. *American Educational Research Journal*, *23*, 558–571.
- Boyd-Zaharias, J., & Pate-Bain, H. (2000). Early and new findings from Tennessee's Project STAR. In M. C. Wang & J. D. Finn (Eds.), *How small classes help teachers do their best* (pp. 65–97). Philadelphia: Temple University Center for Research in Human Development and Education.
- Bray, R. M., Kerr, N. L., & Atkin, R. S. (1978). Effects of group size, problem difficulty, and sex on group performance and member reactions. *Journal of Personality and Social Psychology*, *36*, 1224–1240.
- Bryk, A. S., & Thum, Y. M. (1989). The effects of high school organization on dropping out: An exploratory investigation. *American Educational Research Journal*, *26*, 353–383.
- Cahen, L. S., Filby, N., McCutcheon, G., & Kyle, D. W. (1983). *Class size and instruction*. New York: Longman.
- Caldas, S. J. (1993). Reexamination of input and process factor effects on public school achievement. *Journal of Educational Research*, *86*, 206–214.
- Chase, C. I., Mueller, D. J., & Walden, J. D. (1986). *PRIME TIME: Its impact on instruction and achievement*. Indianapolis: Indiana Department of Education.
- CSR Research Consortium. (2000). *Class size reduction in California: The 1998–99 evaluation findings*. Sacramento: California Department of Education.
- Dennis, B. D. (1986). *Effects of small class size (1:15) on the teacher/learning process in grade two*. Unpublished doctoral dissertation, Tennessee State University, Nashville.
- Egelson, P., & Harman, P. (2000). Ten years of small class size in Burke County, North Carolina. In M. C. Wang & J. D. Finn (Eds.), *How small classes help teachers do their best* (pp. 279–297). Philadelphia: Temple University Center for Research in Human Development and Education.
- Everston, C. M., & Folger, J. K. (1989, March). *Small classes, large classes: What do teachers do differently?* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Finn, J. D., & Achilles, C. M. (1999). Tennessee's class size study: Findings, implications, misconceptions. *Educational Evaluation and Policy Analysis*, *21*, 97–109.
- Finn, J. D., Folger, J., & Cox, D. (1991). Measuring participation among elementary grade students. *Educational and Psychological Measurement*, *51*, 393–402.

- Finn, J. D., Forden, M. A., Verdinelli, S., & Pannozzo, G. M. (2001). *Evaluation of the class size reduction initiative—Buffalo Public Schools*. Buffalo, NY: University at Buffalo, Graduate School of Education.
- Finn, J. D., Fulton, D., Zaharias, J., & Nye, B. A. (1989). Carry-over effects of small classes. *Peabody Journal of Education*, 67, 75–84.
- Finn, J. D., Gerber, S. B., Achilles, C. M., & Boyd-Zaharias, J. (2001). The enduring effects of small classes. *Teachers College Record*, 103, 45–83.
- Finn, J. D., Pannozzo, G. M., & Voelkl, K. E. (1995). Disruptive and inattentive-withdrawn behavior and achievement among fourth graders. *Elementary School Journal*, 95, 421–434.
- Finn, J. D., & Rock, D. M. (1997). Academic success among students at risk for school failure. *Journal of Applied Psychology*, 82, 221–234.
- Finn, J. D., & Wang, M. C. (Eds.). (2002). *Taking small classes one step further*. Greenwich, CT: Information Age.
- Forsyth, D. R. (1999). *Group dynamics* (3rd ed.). Belmont, CA: Brooks-Cole.
- French, J. R. P., Jr. (1941). The disruption and cohesion of groups. *Journal of Abnormal and Social Psychology*, 36, 361–377.
- French, R. L., & Galloway, C. M. (n.d.). *Communication events: A new look at classroom interactions*. Knoxville: University of Tennessee.
- Gibb, J. R. (1951). The effects of group size and of threat reduction upon certainty in a problem-solving situation. *American Psychologist*, 6, 324.
- Glass, G. V., & Smith, M. L. (1978). *Meta-analysis of research on the relationship of class size and achievement*. San Francisco: Far West Laboratory for Educational Research and Development.
- Goodenow, C. (1993). Classroom belonging among early adolescent students. *Journal of Early Adolescence*, 13, 21–43.
- Hanushek, E. A. (1998). *The evidence on class size*. Rochester, NY: University of Rochester, W. Allen Wallis Institute of Political Economy.
- Hardy, C., & Latane, B. (1986). Social loafing on a cheering task. *Social Science*, 71, 165–172.
- Harkins, S. (1987). Social loafing and social facilitation. *Journal of Experimental Social Psychology*, 23, 1–18.
- Harkins, S., & Jackson, J. (1985). The role of evaluation in eliminating social loafing. *Personality and Social Psychology Bulletin*, 11, 457–465.
- Haskins, R., Walden, T., & Ramey, C. T. (1983). Teacher and student behavior in high- and low-ability groups. *Journal of Educational Psychology*, 75, 865–876.
- Hedges, L. V., Laine, R. D., & Greenwald, R. (1994). Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher*, 23, 5–14.
- Hogg, M. A. (1992). *The social psychology of group cohesiveness: From attraction to social identity*. New York: New York University Press.
- Homans, G. C. (1974). *Social behavior: Its elementary forms*. New York: Harcourt Brace Jovanovich.
- Johnston, J. M. (1990). *What are teachers' perceptions of teaching in different classroom contexts?* Paper presented at the annual meeting of the American Educational Research Association, Boston, MA. (ERIC Document Reproduction Service No. ED320867).
- Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology*, 65, 681–706.
- Karau, S. J., & Williams, K. D. (1995). Social loafing: Research findings, implications, and future directions. *Current Directions in Psychological Science*, 5, 134–140.

- Kashti, Y., Arieli, M., & Harel, Y. (1984). Classroom seating as a definition of situation: Observations in an elementary school in one development town. *Urban Education, 19*, 161–181.
- Kerr, N., & Bruun, S. (1981). Ringelmann revisited: Alternative explanations for the social loafing effect. *Personality and Social Psychology Bulletin, 7*, 224–231.
- Kiser-Kling, K. A. (1995). *A comparative study of life in first grade classrooms of 1:14 and 1:23 teacher/pupil ratios*. Unpublished doctoral dissertation, University of North Carolina, Greensboro.
- Krieger, J. D. (2001). *Nature of teacher/student interactions in public elementary schools: Does class size make a difference?* Unpublished doctoral dissertation, University of New Orleans.
- Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics, 114*, 497–532.
- Lamborn, S. D., Brown, B. B., Mounts, N. S., & Steinberg, L. (1992). Putting school in perspective: The influence of family, peers, extracurricular participation, and part-time work on academic engagement. In F. M. Newmann (Ed.), *Student engagement and achievement in American secondary schools* (pp. 153–181). New York: Teachers College Press.
- Latane, B., & Nida, S. (1981). Ten years of research on group size and helping. *Psychological Bulletin, 39*, 308–324.
- Latane, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology, 37*, 823–832.
- Lewit, E. M., & Baker, L. S. (1997). Class size. *Future of Children, 7*(3), 112–121.
- Maier, P., Molnar, A., Percy, S., Smith, P., & Zahorik, J. (1997). *First year results of the Student Achievement Guarantee in Education program*. Milwaukee: University of Wisconsin, Center for Urban Initiatives and Research.
- Marks, H. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal, 37*, 153–184.
- McDermott, P. A., & Beitman, B. S. (1984). Standardization of a scale for the study of children's learning styles: Structure, stability, and criterion validity. *Psychology in the Schools, 21*, 5–14.
- McMillan, D. W., & Chavis, D. M. (1986). Sense of community: A definition and theory. *Journal of Community Psychology, 14*, 6–23.
- Miles, K. H. (1995). Freeing resources for improving schools: A case study of teacher allocation in Boston Public Schools. *Educational Evaluation and Policy Analysis, 17*, 476–493.
- Molnar, A., Smith, P., & Zahorik, J. (1999). *1998–99 evaluation results of the Student Achievement Guarantee in Education (SAGE) program*. Milwaukee: University of Wisconsin, School of Education.
- Molnar, A., Smith, P., & Zahorik, J. (2000). *1999–2000 evaluation results of the Student Achievement Guarantee in Education (SAGE) program*. Milwaukee: University of Wisconsin, School of Education.
- Mosteller, F. (1995). The Tennessee study of class size in the early school grades. *Future of Children, 5*, 113–127.
- Mullen, B., & Copper, C. (1994). The relation between group cohesiveness and performance: An integration. *Psychological Bulletin, 115*, 210–237.
- Northwest Regional Educational Laboratory. (1984). *Effective schooling practices: A research synthesis*. Portland, OR: Author.
- Osterman, K. F. (2000). Students' need for belonging in the school community. *Review of Educational Research, 70*, 323–367.

- Peterson, P. L., Swing, S. R., Stark, K. D., & Waas, G. A. (1984). Students, cognitions and time on task during mathematics instruction. *American Educational Research Journal*, 21, 487–515.
- Petty, R., Harkins, S., Williams, K., & Latane, B. (1977). The effects of group size on cognitive effort and evaluation. *Personality and Social Psychology Bulletin*, 3, 579–582.
- Prout, J. M. (2000). *Indoor air quality: Multiple implications of CO₂ in the classroom environment*. Unpublished doctoral dissertation, Eastern Michigan University.
- Ringelmann, M. (1913). Recherches sur les moteurs animés: Travail de l'homme. *Annales de l'institut National Agronomique*, 12, 1–40.
- Robinson, G. E. (1990). Synthesis of research on effects of class size. *Educational Leadership*, 47(7), 80–90.
- Royal, M. A., & Rossi, R. J. (1996). Individual-level correlates of sense of community: Findings from workplace and school. *Journal of Community Psychology*, 24, 395–416.
- Sanna, L. J. (1992). Self-efficacy theory: Implications for social facilitation and social loafing. *Journal of Personality and Social Psychology*, 62, 774–786.
- Sarason, S. B. (1974). *The psychological sense of community: Prospects for a community psychology*. San Francisco: Jossey-Bass.
- Shapson, S. M., Wright, E. N., Eason, G., & Fitzgerald, J. (1980). An experimental study of the effects of class size. *American Educational Research Journal*, 17, 141–152.
- Shaw, M. E. (1976). *Group dynamics: The psychology of small group behavior* (2nd ed.). New York: McGraw-Hill.
- Smith, L., Ross, S., Alberg, M., & Lowther, D. (1999). *School observation measure: Observer's manual*. Memphis, TN: University of Memphis, Center for Research in Educational Policy.
- Spivak, G., & Cianci, N. (1987). High-risk early behavior pattern and later delinquency. In J. D. Burchard & S. N. Burchard (Eds.), *Prevention of delinquent behavior* (pp. 44–74). Beverly Hills, CA: Sage.
- Stasz, C., & Stecher, B. (2002). Before and after class-size reduction: A tale of two teachers. In M. C. Wang & J. D. Finn (Eds.), *Taking small classes one step further* (pp. 19–50). Greenwich, CT: Information Age.
- Swift, M. S., & Spivak, G. (1968). The assessment of achievement-related classroom behavior. *Journal of Special Education*, 2, 137–149.
- Szymanski, K., & Harkins, S. G. (1993). The effect of experimenter evaluation on self-evaluation within the social loafing paradigm. *Journal of Experimental Social Psychology*, 29, 268–286.
- Veldman, D. J., & Worsham, M. (1983). Types of student classroom behavior. *Journal of Educational Research*, 76, 204–209.
- Voelkl, K. E. (1995). *Class size and classroom participation*. Paper presented at the National Symposium on Class Size Research, London, England.
- Wang, M. C., & Stull, J. C. (2000). School characteristics and classroom practice: Smaller versus larger classes. In M. C. Wang & J. D. Finn (Eds.), *How small classes help teachers do their best* (pp. 175–198). Philadelphia: Temple University Center for Research in Human Development and Education.
- Wehrlage, G. G., Rutter, R. A., Smith, G. A., Leski, N., & Fernandez, R. R. (1989). *Reducing the risk: Schools as communities of support*. Philadelphia: Falmer Press.
- Wolff, K. H. (1950). *The sociology of Georg Simmel*. New York: Free Press.
- Word, E., Johnston, J., Bain, H., Fulton, D. B., Boyd-Zaharias, J., Lintz, M. N., et al. (1990). *Student/Teacher Achievement Ratio (STAR): Tennessee's K–3 class size study*. Nashville: Tennessee State Department of Education.

Wright, E. N., Shapson, S. M., Eason, G., & Fitzgerald, J. (1977). *Effects of class size in the junior grades: A study*. Toronto: Ontario Ministry of Education.

Authors

JEREMY D. FINN is a Professor of Education, University at Buffalo, State University of New York, Graduate School of Education, 409 Christopher Baldy Hall, Buffalo, NY 14260; e-mail finn@buffalo.edu. His research interests include multivariate statistical methods and studies of students at risk and educational equity.

GINA M. PANNOZZO is a graduate assistant, University at Buffalo, State University of New York, Graduate School of Education, 409 Christopher Baldy Hall, Buffalo, NY 14260; e-mail pannozzo@buffalo.edu. Her research interests include class size and student engagement and the effects of classroom culture and school community on teaching and learning.

CHARLES M. ACHILLES is a Professor of Education Administration, Seton Hall University and Eastern Michigan University. He can be contacted at 4477 Snug Harbor, Geneva, NY 14456; e-mail plato936@rochester.rr.com. His research interests include class size and student outcomes, school organization, and education administration.

APPENDIX A
Studies Reviewed

Venue	Authors	Subjects	Method and measures ^a	Focus
High Point, North Carolina	Achilles, Kiser-Kling, Aust, & Owen (1995); Kiser-Kling (1995)	Grades K-2 in 2 Title I schools with a focus on Grade 1 4 small Grade 1 classes 2 traditional Grade 1 classes Small class enrollment = 14 Traditional class enrollment = 23	Observational study Structured classroom observations using PIT method (French & Galloway, n.d.) Interactions among students and between teachers and students coded every 4-5 seconds Discipline referral records kept by school principal (Kiser-Kling: interviews with 4 Grade 1 teachers of small classes)	Learning behavior Antisocial behavior Teacher interpersonal
England	Blatchford (in press)	"Reception classes" (ages 4-5) 235 students 39 classes 21 small classes 18 large classes Small class enrollment < 20 Large class enrollment > 30	Observational study Formal classroom observations Observed 6 students in each class in 10-second intervals for 5-minute periods, 5 times per day Pupil Behaviour Ratings (PBR) Scale	Learning behavior Antisocial behavior Prosocial behavior Teacher interpersonal

Australia	Bourke (1986)	Grade 5 33 schools 63 teachers Class size ranged from 12 to 33	Observational/survey Formal classroom observations Class size was averaged from ratio of no. of students to no. of teachers present at 5-minute intervals Observers recorded teaching practices used over 12-week period	Teacher interpersonal
STAR Tennessee	Boyd-Zaharias & Pate-Bain (2000)	Grade 1: 49 teachers Grades 2 and 3: 130 teachers 65 high quality (top 10%) 65 low quality (bottom 50%) Small class enrollment = 15-17 Regular class enrollment = 22-26	Observational/interview Formal classroom observations Teachers were rated as <i>poor</i> , <i>fair</i> , <i>good</i> , or <i>excellent</i> on a number of criteria adapted from the Concepts of Effective Teaching (NWREL, 1984) Information about length of observations or how ratings were derived was not provided	Teacher interpersonal
California and Virginia	Cahen, Filby, McCutcheon, & Kyle (1983)	Grade 2 classes 4 classes in two schools—1 in Virginia, 1 in California Virginia enrollment = 20, then reduced to 13 midway through year California Enrollment = 30, then reduced to 22 midway through year	Observational study Formal classroom observations Modified Flanders system used to observe 6 target students during reading and math	Learning behavior

(continued)

APPENDIX A (Continued)

Venue	Authors	Subjects	Method and measures ^a	Focus
Prime Time Indiana	Chase, Mueller, & Walden (1986)	<p>Grades 1 and 2 137 teachers 29 schools Small class enrollment = 18 Classes with teacher aide enrollment = 24</p>	<p>Survey and observational study Teacher reports (surveys)</p>	<p>Learning behavior Antisocial behavior Prosocial behavior</p>
California	CSR Research Consortium (2000)	<p>Grade 3 636 teachers 531 teachers in reduced classes 105 teachers in nonreduced classes Reduced class enrollment = 20 or less Nonreduced class enrollment was not provided</p>	<p>Survey State-wide teacher survey</p>	<p>Learning behavior Antisocial behavior Prosocial behavior</p>
Tennessee	Dennis (1986)	<p>Grade 2 255 students Experimental group = 6 small classes in one school Control group = 4 regular classes in a matched school Blind control group = 85 students from metropolitan Nashville schools Small class enrollment = 13-15 Control class enrollment = 25</p>	<p>Survey Corporal punishment records for students from first 6 months of school year when small classes were instituted</p>	<p>Antisocial behavior</p>

Burke County, North Carolina	Egelson & Harman (2000)	Grades 1-3 5 schools 8-10 small classes in each school Small class enrollment = 18	Observational Unannounced observations, 15 minutes in length School observation measure	Learning behavior
STAR Tennessee	Everston & Folger (1989)	Grade 2 13 schools 52 classes Small enrollment = 13-17 Regular enrollment = 21-25 Regular/aide = 21-25	Observational Observed in fall and again in winter during math and read- ing lessons Recorded count of interactions between teachers and students as behavioral, procedural, or academic	Learning behavior Antisocial behavior
Buffalo, New York	Finn, Forden, Verdinelli, & Panno (2001)	Grades K-3 213 teachers Small class enrollment = 20 or less Other class configuration enroll- ment = 21-25	Survey/teacher report Questionnaire administered June 2001	Antisocial behavior Prosocial behavior Teacher interpersonal
STAR Tennessee	Finn, Fulton, Zaharias, & Nye (1989)	Grade 4 58 schools 2,207 students Small class enrollment = 13-17 Regular and Regular with teacher aide class enrollment = 22-26	Survey/teacher report Student Participation Questionnaire Teachers completed question- naire for no more than 10 stu- dents in each class for only those students who had been in same class type for Grades 1-3	Learning behavior Antisocial behavior

(continued)

APPENDIX A (Continued)

Venue	Authors	Subjects	Method and measures ^a	Focus
STAR Tennessee	Johnston (1990)	Grades K-3 1,000 teachers Small class enrollment = 13-17 Regular and Regular with teacher aide class enrollment = 21-25	Qualitative-interview Exit interviews during May of each year of project Kindergarten interviews used to identify and define themes 14 themes used to categorize interview protocols for remaining interviews	Antisocial behavior Prosocial behavior Teacher interpersonal
Louisiana	Krieger (2001)	Grades K-2 11 teachers in 3 schools 256 students 4 small classes with enrollment < 18 7 large classes with enrollment > 24	Qualitative-ethnography Classes were videotaped during language arts and mathematics lessons in spring 1999 Videotapes were analyzed using comparative analysis to char- acterize the verbal and non- verbal interactions between teachers and students Interactions also coded using PIT method (French & Galloway, n.d.)	Teacher interpersonal
SAGE Wisconsin	Maier, Molnar, Percy, Smith, & Zahorik (1997)	Grades K and 1 ~166 teachers ~6,000 students Small class enrollment = 15 Alternative models—team, shared space, floating teacher, split day with various enrollments	Survey/teacher report Modified Student Participation Questionnaire, completed fall and spring	Learning behavior

SAGE Wisconsin	Molnar, Smith, & Zahorik (1999): 2nd- and 3rd-year executive summary	Grades K-3 Small enrollment = 15 Alternative models—team, shared space, floating teacher, split day	Qualitative—case study Case studies of 3 classrooms, Grades 1, 2, and 3, in each of 3 schools—9 classes that included classroom observations, and teacher interviews	Antisocial behavior Prosocial behavior Teacher interpersonal
Toronto, Ontario, Canada	Shapson, Wright, Eason, & Fitzgerald (1980)	Grades 4 and 5 11 schools 62 classes Class sizes of 16, 23, 30, 47	Observational and survey Toronto Classroom Observation Scale (TCOS) 8 half-day observations/class/year Teacher and student reports Teachers in class of 16 or 23 in first year were placed in class of 30 or 37 in second Students in class of 16 or 37 in first year were placed in class of different size in second	Learning behavior
California	Stasz & Stecher (2002)	Grade 3 16 teachers over 2 years	Qualitative—case study Classroom observations and videotaped lessons during language arts and mathematics Teacher interviews Principal interviews Content surveys Daily activity logs Student work samples	Teacher interpersonal

(continued)

APPENDIX A (Continued)

Venue	Authors	Subjects	Method and measures ^a	Focus
STAR Tennessee	Voelkl (1995)	Grade 8 1,481 students who had been involved with Project STAR in Grades K-3	Survey/teacher report Student Participation Questionnaire Math or English teachers com- pleted questionnaire for no more than 10 students in each class for students who had been in same class type for Grades 1-3	Learning behavior

^aThe methods and measures listed are only those discussed in this article and do not reflect the full scope of those studied.

APPENDIX B

Major Class-Size Reduction Initiatives

Project Prime Time: Indiana

Between 1984 and 1986, Indiana launched Project Prime Time as a statewide initiative to reduce class size in elementary classes (Grades 1–3) to 18, or 24 with an instructional assistant. Grade 1 classes were reduced in 1984, followed by Grade 2 classes in 1985 and Grade 3 classes in 1986. An evaluation of the effectiveness of Prime Time was completed using data collected from 29 of the 302 school districts in the state. Information was gathered via classroom observations (300 classes in Grades 1 and 2), teacher surveys (349 in Grades 1 and 2; 326 in Grades 3 and 4), parent surveys (1,252 in Grades 1 and 2), and interviews with teachers ($N = 167$), parents ($N = 275$), and principals ($N = 59$). In addition, achievement data from standardized achievement tests given in 11 districts were analyzed. The evaluation reported mixed results when pre and post achievement test scores in mathematics and reading were compared. Overall, a higher percentage of students in small classes showed a significant gain in Grade 1 than in Grade 2, and gains in reading were larger than those in mathematics (Chase, Mueller, & Walden, 1986).

Project STAR: Tennessee

From 1985–1989, the STAR (Student Teacher Achievement Ratio) experiment was conducted in Tennessee. Project STAR was a statewide randomized experiment involving 76 schools, over 1,200 teachers, and almost 12,000 students over 4 years. Students entering kindergarten were assigned at random to a small class (13–17 students), a regular class (22–26 students), or a regular class with a full-time teacher aide within each participating school. Students were kept in the same class arrangement for up to 4 years, through Grade 3. Teachers were also randomly assigned to classes, with a new teacher assigned to the class in each successive grade. Data were collected on a broad range of measures in each of the 4 years of the study via classroom observations, teacher interviews, and teacher questionnaires. All pupils returned to full-size classes in Grade 4. Follow-up achievement test data were collected in Grades 4–8. In addition, teachers rated students' classroom behavior in Grades 4 and 8, and students' attitudes toward school were assessed in Grade 8. Small classes were found to have significant advantages in all subject areas both during the period of small classes (K–3) and in every subsequent grade (Finn & Achilles, 1999; Finn, Gerber, Achilles, & Boyd-Zaharias, 2001; Word et al., 1990). Effect sizes for minority students in K–3 were typically two to three times as large as those for White students.

Burke County, North Carolina

The Burke County, North Carolina, class-size reduction initiative, begun in 1991, involved a total of 14 elementary schools. With the goal of improving education in this high-poverty community, a pilot program in 1991–1992 reduced class size to 15 in Grade 1 in four schools, and in Grades 2 and 3 in subsequent years. On the strength of positive pilot program results, the program was expanded to include all Grade 1 and 2 classes in all 14 elementary schools, and Grade 3 in six of the schools. By reallocating existing resources, the Burke County initiative was able to implement small class sizes with no increase in per-pupil expenditures for the district. Data for the evaluation of the initiative's effectiveness were collected via questionnaires and surveys of 28 Grade 1 teachers in 1991–1992 and 112 Grade 1–3 teachers in 1993–1994, classroom observations conducted in spring of 1995 and October 1999, and interviews with teachers, principals, and parents. In addition, end-of-year achievement test scores were collected in Grades 1–3. Results showed that small classes outperformed a comparison group of regular classes in mathematics and reading at the end of Grades 1, 2, and 3. In addition, after returning to regular classes, students in small classes continued to outperform other students in Grades 4 and 7 (Egelson & Harman, 2000).

Project SAGE: Wisconsin

Wisconsin's Student Achievement Guarantee in Education (SAGE) Program, begun in 1996, was targeted to low-income schools (i.e., schools in which 50% of the enrollment qualified for Title 1 services). Funding was provided up to \$2,000 per student enrolled in SAGE classrooms (K–3). 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 and collaborate with community organizations to provide educational, recreation, and community and social services; provide a rigorous academic curriculum to improve academic achievement; and establish staff development and accountability mechanisms.” During 1996–1997, 30 schools in 21 districts, including 7 in Milwaukee, began the program in K–1. Grade 2 was added in these schools in 1997–1998, and Grade 3 in 1998–1999. To date, 264 self-contained small classes (average class size: 14 students) and 92 “alternative model” classes in Grades K–3 are involved in SAGE. Alternative models included classes that were team taught with two full-time teachers taking responsibility for approximately 30 students; “shared space,” where a classroom has been fitted with a temporary wall to create two separate teaching spaces each accommodating one teacher and 15 students; and “floating teacher,” with one full-time teacher assigned to a class of 30 and a rotating teacher who joins the class during some lessons. The SAGE evaluation used matched comparison classrooms from the same districts that were not reduced in size. Data were collected via teacher and principal interviews, teacher questionnaires, and classroom observations during each year of the project. Small classes were found to have a significant impact on student achievement in all grades, with the greatest achievement gains in Grade 1. As in Project STAR, the results for minority students were especially pronounced (Molnar, Smith, & Zahorik, 1999, 2000).

Statewide Class-Size Reduction: California

In 1996, the California state legislature passed a reform initiative to limit class size statewide in Grades K–3 to a maximum of 20 students. Within a period of several months, new teachers were hired and placed in “small” Grade K–3 classrooms across the state. In 3 years of operation, this largest CSR initiative has resulted in the hiring of approximately 28,000 new teachers and virtually every classroom in Grades 1–2 being reduced in size. As of 1998–1999, small classes had been implemented in 86% of all kindergarten classes, 100% of all Grade 1 classes, 98% of all Grade 2 classes, and 84% of all Grade 3 classes. Data for the evaluation were collected from a number of sources: teacher questionnaires, teacher interviews, classroom observations, and the California Basic Educational Data System. Because the program was implemented so quickly, few large classes were available to serve as a comparison group. The evaluation of academic achievement has focused on Grade 3, in which small but statistically significant gains were reported in reading and mathematics (CSR Research Consortium, 2000). The California evaluation also examined other features of the program, among them the negative impact of hiring many new teachers in a short time span. In addition, several studies of classroom processes were undertaken. Case studies focused on changes in teachers' practice as they switched from teaching regular classes to teaching small classes.

A vertical bar on the left side of the page, consisting of a series of horizontal segments in shades of yellow and orange, with a small red diamond at the top.

COPYRIGHT INFORMATION

TITLE: The "Why's" of Class Size: Student Behavior in Small
Classes

SOURCE: Rev Educ Res 73 no3 Fall 2003

WN: 0328800924002

The magazine publisher is the copyright holder of this article and it is reproduced with permission. Further reproduction of this article in violation of the copyright is prohibited. To contact the publisher:
<http://www.aera.net/>

Copyright 1982-2003 The H.W. Wilson Company. All rights reserved.