

## **School Size in Chicago Elementary Schools: Effects on Teachers' Attitudes and Students' Achievement**

**Valerie E. Lee**  
*University of Michigan*  
**Susanna Loeb**  
*Stanford University*

*This study explores whether teachers and students are influenced by the size of the inner-city elementary school to which they belong. Focusing on teachers' attitudes about their responsibility for student learning and students' 1-year gains in mathematics achievement scores, we used data from almost 5,000 teachers and 23,000 sixth and eighth-grade students in 264 K-8 Chicago schools. The data were collected through 1997 surveys and annual standardized tests. We employed hierarchical linear modeling (HLM) to estimate school effects. On both outcomes, small schools (enrolling fewer than 400 students) are favored compared with medium-sized or larger schools. In small schools, teachers have a more positive attitude about their responsibility for students' learning and students learn more. Even after taking size into account, learning is also higher in schools with higher levels of collective responsibility. Thus, we conclude that school size influences student achievement directly and indirectly, through its effect on teachers' attitudes.*

---

VALERIE E. LEE is Professor, School of Education, University of Michigan, 601 S. University Ave., Ann Arbor, MI 48109. Her specializations are sociology of education and qualitative research methods.

SUSANNA LOEB is Assistant Professor, School of Education, Stanford University, CERAS, Stanford, CA 94304. Her specialization is economics of education.

---

Increasingly, educational researchers and practitioners agree that the structural characteristics of schools have an important impact on the lives of school members. A crucial element of any school's structure is the number of students enrolled, that is, the size of the school. Almost all empirical studies that estimate the effects of school size have been conducted in high schools. They have also focused almost exclusively on how school size influences *students*.

This study expands the body of research on school size in several ways. First, we consider physical location. We investigate the impact of school size on schools within a single large school district (Chicago), which reduces the empirical difficulties of controlling for cross-district differences. A second (and related) innovation is the social composition of the schools studied. Chicago's public schools enroll high proportions of socially disadvantaged students. Lee and Smith (1997) highlighted the special salience of size in schools like those in Chicago, which have a high concentration of minority and low-income students. Third, we investigate the effects of school size in elementary schools. All the schools in this study include the full range of elementary and middle-school grade levels (K–8); other studies of school size included only high schools. A fourth difference is a broader set of school members and outcomes that might be influenced by school size. We investigate potential size effects on all teachers in each school (regardless of grade level or subject specialty) and on all students in Grades 6 and 8. In particular, we explore size effects on student achievement and on teachers' attitudes about their students. We hypothesize that school size influences student achievement, both directly and indirectly, through teachers' attitudes.

Beyond these differences, our study shares several design characteristics with a study by Lee and Smith (1997), who investigated high school size. Both studies have a longitudinal design (we assess achievement growth in mathematics scores over one school year, 1996–1997). Both studies use large samples. Rather than representative samples, however, the participants in this study represent the population. Nearly all teachers and students in Grades 6 and 8 in Chicago elementary (K–8) schools completed the surveys from which these data were drawn. Our sample includes 264 Chicago elementary schools, 4,495 teachers, and 22,599 sixth and eighth graders in the same schools. Both studies use the same methodological approach, hierarchical linear modeling, (HLM; Bryk & Raudenbush, 1992).

## Background

### Two Distinct Research Bases

The research background for this study lies in two distinct and nonoverlapping bodies of literature: writings about school size and teachers' work. Unfortunately, the major focus of the study, school size, is supported by a body of research that almost entirely targets high schools. Moreover, the empirical research on school size generally targets its effects on students.

## *Effects of School Size on Attitudes and Achievement*

Because we examine how size affects teachers as well as students, another body of research is relevant, one that focuses on teachers' work (again, conducted mostly in secondary schools). The thread with which we combine these two bodies of literature is a focus on the organizational features of schools.

### **Research on School Size**

*Conflicting research streams and conclusions.* Enrollment size is an important ecological feature of any educational organization. There are two relevant streams of research, one of which is sociological in nature. This stream examines how size influences other organizational properties. As schools grow, they become more bureaucratic. Consequences flow from such changes, including more specialized instructional programs and more formal human relationships. The other stream, with an economic bent, investigates the savings that might accrue from expanding the scale. This research highlights the potential for increased efficiency through reduced redundancy and increased resource strength as schools get bigger (Buzacott, 1982). These arguments have driven the school consolidation movement (Kenny, 1982). Although the rationale is that greater size results in a more efficient operation (Guthrie, 1979; Michelson, 1972), savings projected by proponents of school consolidation have not materialized. As schools get larger, they usually expand their administrative staffs. In rural areas, increased costs for transportation and distribution of school materials often offset any savings (Chambers, 1981).

Conclusions from these two streams do not always coincide. Although the studies with an organizational focus generally favor small schools, the economic research suggests that there may be some benefits from increased size. Organizational research on school size also separates into two branches; studies focus on either the schools' academic or social organization. The latter branch, which is most relevant here, depends on the former.

*Size and specialization.* Research documents a relationship between organizational size and program specialization. In principle, large schools have more students with similar needs and are better able to create specialized programs to address such needs (again, the efficiency argument). In contrast, small schools must direct their resources to core programs, with marginal students (at either end of a distribution of ability or interest) either excluded from programs or absorbed into programs that may not meet their needs (Monk, 1987; Monk & Haller, 1993). Despite the logic of the specialization argument, research on high school tracking suggests that extensive differentiation in curricular offerings and students' academic experiences may have debilitating consequences (Gamoran, 1989; Lee & Bryk, 1989; Oakes, 1985). Increasing size invites curriculum specialization, resulting in differentiation of academic experiences and social stratification of outcomes (Bryk, Lee, & Holland, 1993; Lee & Bryk, 1989; Lee, Bryk, & Smith, 1993).

Although both the specialization and efficiency arguments are most relevant for secondary schools, in theory these arguments may apply to all schools.

*Size, social equity, and learning.* Although much extant research is framed within a bigger versus smaller mode, the objective of one study was to estimate an appropriate balance point for learning as a function of high school size (Lee & Smith, 1997). Achievement gains in mathematics and reading in high school were largest in schools with 600–900 students. Schools that were smaller or larger than this had lower average gains, with schools enrolling more than 2,000 students being the most disadvantaged. Small schools were favored in terms of social equity; the relationship between social class and achievement gains was lower in schools enrolling less than 900 students. It was also reported in this study that even though the same ‘ideal size’ was consistent across schools identified by their average socioeconomic status (SES) and minority concentration, school size was a more important factor in determining learning in schools enrolling many economically disadvantaged and/or minority students. There has been no similar study of school size in elementary schools.

*Size and the communitarian argument.* Research on the communal aspects of learning is critical of program specialization, in terms of both the curriculum and other student experiences in school. Sociological theory suggests that human interactions and ties become more formal as organizations grow (Weber, 1947). Organizational growth generates new bureaucratic structures, as connections between individuals become less personal. These structures inhibit communal organization (Bryk & Driscoll, 1988; Bryk et al., 1993). The theory has been confirmed in research identifying organizational characteristics of effective schools. In school climate studies, for example, size serves as an ecological feature of the social structure, part of a physical environment influencing social interactions (Barker & Gump, 1964; Bryk & Driscoll, 1988; Garbarino, 1980).

This line of research concludes that small schools benefit students in several ways. The more constrained curriculum in many small high schools includes academic courses, so that almost all students have similar academic experiences regardless of their interests, abilities, or social background. This results in both higher average achievement and achievement that is more equitably distributed (Lee & Bryk, 1988, 1989). Social relations are also more positive in small schools. The preponderance of sociological evidence on high schools suggests that except in the extreme, smaller is better (Lee et al., 1993).

### Relevant Research on Teachers’ Work

*Collaboration, cooperation, and norms.* Calls for organizational change in schools have included reforms targeting teachers and teaching (Carnegie Task Force on Teaching as a Profession, 1986; National Commission on Excellence in Education, 1983). Such calls are consistent with research and writing that focus on school structural features that are related to

teachers' work (Johnson, 1990; Lieberman, 1992; Rosenholtz, 1989; Rowan, 1990a). An underlying assumption in reform proposals that target teachers' work is its link, implicit or explicit, to student learning.

Relevant research focuses on the social dimensions of teachers' work: sharing, collaboration, and a normative school culture defined by teachers' attitudes (Firestone & Herriot, 1982; Firestone & Wilson, 1985; Lieberman, 1992; Louis & Kruse, 1995; Louis, Marks, & Kruse, 1996). The solitary nature of teaching can be ameliorated through face-to-face relationships with colleagues and by leaders who support collegiality (Lortie, 1975). However, developing a professional community in a school must work against an organizational culture that typically centers more on specialization than on cooperation. Besides the obvious psychic benefits that accrue to teachers through social contact with their peers, cooperative professional relationships are also important in developing a productive school culture (Little, 1982, 1990). Collegial interactions foster the sharing of information and advice about teaching (Rowan, 1990b).

Individual commitment to organizational goals develops in response to how people use power, often exercised through decision-making strategies and sanctions (i.e., rewards and/or sanctions) held for normative activity (Etzioni, 1975). Commitment is part of an individual's positive response to the exercise of power. However, in schools, the normative value typically placed on objectivity and fairness also emphasizes the role of expert authority (Greene & Podsakoff, 1981). Teachers' commitment to the goals and values of their schools, in the aggregate, represents a feature of a school's culture (Wyner, 1991). However, a positive school professional community requires commitment in a particular direction: a focus on student learning and engagement in academic pursuits of all students (Louis et al., 1996). Teachers provide students with more than information to be digested and procedures to be mastered. They also communicate their own attitudes toward particular students and toward the learning process (Lee & Smith, 1996). Teachers' attitudes are instrumental in how students construct knowledge, which ultimately results in learning (Brophy, 1986).

*Effects of teachers' attitudes on students.* When it first appeared, research about the self-fulfilling prophecy of teachers' expectations on student learning (Rosenthal & Jacobson, 1968) provoked much controversy. However, this link is by now well documented and widely accepted (Brophy, 1983; Cooper, Findley, & Good, 1982; Cooper & Tom, 1984; Raudenbush, 1984). The conclusion is consistent: teachers' beliefs about students' ability to learn influence achievement. Not only are students more engaged in the learning process when teachers have high expectations for them and are willing to take personal responsibility for them (Firestone & Rosenblum, 1988), they also learn more (Cooper & Tom, 1984).

Most research on teacher expectations has been conducted with a psychological orientation, which typically matches individual students with their teachers. However, a recent study expanded the notion of teachers' expectations in a sociological direction through the idea of teachers' *collective*

responsibility for student learning (Lee & Smith, 1996). This construct might describe schools on a continuum based on average teacher attitudes. On one end of the spectrum, we would find schools where most teachers take personal responsibility for the success or failure of their own teaching. Such teachers would see teaching and learning as an interactive process, with students cast as active participants, rather than as a one-way flow of information from teachers to students (Brattesani, Weinstein, & Marshall, 1984). On the other end of the spectrum, we would find schools where most teachers see potential impediments between their own teaching and students' learning, namely, students' ability (or lack of it), students' family background, or their motivation. If students do not learn, these teachers would tend to locate the blame for low performance outside of themselves and their own teaching. This same spectrum of attitudes could apply among individual teachers in the same school, as well as to the average teachers' attitudes across schools.

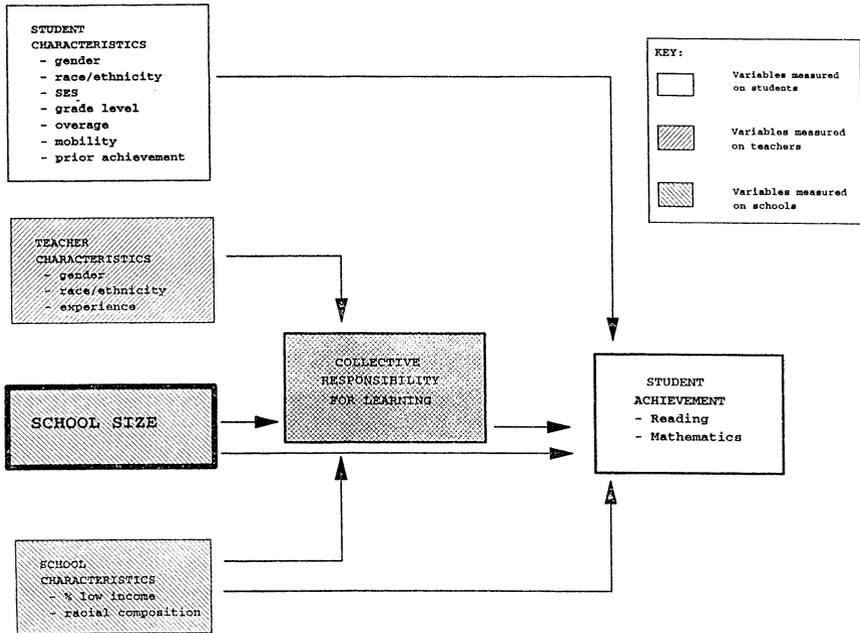
*Collective responsibility for student learning.* This construct may be seen both as a *function* of school organizational structure (i.e., individual teachers' attitudes) or in the aggregate, as an organizational property in itself (i.e., a focus on the collective). Lee and Smith (1996) took the second approach. They found that students showed substantial gains in achievement in four subjects over the first 2 years of high school in schools characterized by (a) higher levels of collective responsibility and (b) more consistency among teachers in the same school in these attitudes. Moreover, they found that learning was distributed more equitably by students' social class in schools with higher levels of collective responsibility.

This study focuses on the construct of collective responsibility for students' learning. We investigate the construct both as a function of school organizational structure (especially school size) and as an organizational property of schools that may directly influence learning. The empirical analysis focuses on the interplay among school size, collective responsibility, and learning.

## **Research Questions**

We organize our investigation around three related research questions. Each successive question requires a positive response to the previous one. Although our research to address these questions is located in the context of elementary schools in a single school district (Chicago), the questions themselves do not depend on this setting. Within each, we pose hypotheses about the effects we expect, including their direction.

Question 1 concerns school size and teachers' attitudes: Is school size related to teachers' assessments of their colleagues' willingness to assume responsibility for students' academic and social development, once demographic characteristics of teachers and schools are taken into account? We hypothesize a negative relationship, that is, teachers in small schools have a higher degree of shared commitment to improve student learning.



**Figure 1. Heuristic model relating school size and collective responsibility for learning to student achievement in Chicago elementary schools**

Question 2 concerns teachers' attitudes and student learning: Is school collective responsibility related to student achievement? We hypothesize a positive relationship, i.e., schools with higher average collective responsibility will also have higher achievement, even after taking into account the social and academic characteristics of students and schools.

Question 3 concerns school size and student learning: Does school size have a direct effect on student achievement, once school collective responsibility and student composition are taken into account? Is the major influence of school size on achievement through teachers' attitudes about collective responsibility (an indirect effect)? We hypothesize that school size has both a direct and an indirect effect on student achievement and that both effects are negative (i.e., size effects favor small schools).

### Conceptual Model

*A school effects study.* Our analyses to test these three research questions are organized around a conceptual model (Figure 1). Both the model and the questions fall within a type of inquiry called school effects research. Studies of this type investigate how characteristics of schools (in this case, enrollment size) influence school members' attitudes and behaviors. Be-

cause we investigate school effects on teachers and students, we formulate a multilevel model, wherein both teachers and students are “nested” in schools. School size, the independent construct, is at the left of the model. This placement suggests that school size is exogenous to the relationships under study (as are teacher, student, and school characteristics). The importance of school size in this model is indicated by the heavy line surrounding its box.

*Two related analyses.* We estimate school effects in two separate but related multilevel analyses, which are organized around the conceptual model represented in Figure 1. We investigate school effects first on teachers and then on students. The first analysis focuses on how school characteristics (size and demographic composition) influence teachers’ attitudes. We target a particular type of attitude, collective responsibility for student learning. Because collective responsibility may be related to teachers’ demographic and professional characteristics (e.g., their own gender, race/ethnicity, and teaching experience), we take these factors into account in the analysis. The particular focus is on how school size influences these attitudes. However, size effects on teachers’ attitudes may also be associated with the types of students attending the schools. Therefore, we also take school demographic composition into account. This first analysis is intended to address Question 1.

The second analysis in Figure 1 focuses on students. Specifically, we target two important school effects on achievement, school social organization (Question 2) and school size (Question 3). As the effects of school social organization and school size on achievement are potentially influenced by school demographic composition, we take this into account. In this case, we conceptualize school social organization as the school average of teachers’ attitudes about collective responsibility. Achievement in this analysis actually measures young adolescents’ learning over the course of a single school year. Therefore, we take into account students’ achievement in the same subject near the end of the previous year.<sup>1</sup> In addition, we expect that learning is influenced by students’ background characteristics. For example, demographics (gender, SES, race/ethnicity) and other factors have been shown to be associated with learning in urban schools (e.g., mobility, whether the student is older than his/her grade-level classmates).

Constructs included in the multilevel conceptual model displayed in Figure 1 are measured at three levels: among teachers, among schools, and among students.<sup>2</sup> The different shadings in the boxes are meant to indicate the different levels of measurement. Collective responsibility is measured at two levels: (1) among teachers for the first analysis and (2) an aggregate organizational characteristic of schools in the second. Its double shading indicates the two levels of measurement.

## **Method**

### **Sample and Data**

The data used in this study were provided by the Consortium on Chicago

## *Effects of School Size on Attitudes and Achievement*

School Research. Organized in 1990, the Consortium is “an independent affiliation of researchers from universities, advocacy groups, and the school system . . . [that conducts] studies to examine the implicit theory behind decentralization, and how it was actually unfolding in elementary schools” (Bryk, Sebring, Kerbow, Rollow, & Easton, 1998, p. xv). Starting in 1991, the Consortium has conducted ongoing surveys in Chicago public schools, organized in part to assess the impact of the Chicago School Reform Act of 1988; Chicago’s teachers, principals, and students are surveyed periodically.

Recent surveys, from which we draw our data, were conducted in the winter of 1997, when the Consortium surveyed 6th, 8th, and 10th-grade students, all teachers, and all principals in the Chicago public schools (Bilcer, 1997). The 1997 surveys were a combined effort of the Consortium and the Chicago Annenberg Study Project. The latter is a multiyear program to document and study the activities of the Chicago Annenberg Challenge. The Chicago Challenge was organized to manage and distribute the funds from a generous 5-year grant from the National Annenberg Challenge to Chicago public schools. Our data are drawn from the 1997 surveys of teachers and students, as well as from standardized test scores from the annual assessments of all Chicago students in all elementary grades.

Our sample is restricted to the 264 Chicago elementary schools that enroll students through the entire span of the elementary and middle grades (K-8), as size is a function of the grades enrolled. Thus, we eliminated a few schools that did not offer the full grade range. The survey response rate for elementary schools was high (88%). Response rates for students and teachers somewhat lower: sixth and eighth-grade students, 78%; elementary teachers, 63% (Bilcer, 1997, p.16). Another sample restriction was imposed; we included only schools that also had data from teachers. Within the 264 schools, our study sample includes 4,495 teachers (an average of 17 teachers per school) and 22,599 students (an average of 86 students per school).

### **Measures**

*Schools.* The major independent variable in this study is school size—the enrollment in each school in 1996–1997. Typical for school size is a positively skewed distribution. This is the case for the enrollment size of Chicago elementary schools (see Figure 2). There are a few very large elementary schools (three schools with 1,500 students or more), several very small schools (24 schools enrolling 350 students or less), with the modal school size around 500 students. Because of this skew, we divided the schools into three size categories: schools enrolling fewer than 400 students ( $n = 25$ ), schools enrolling between 400 and 750 students ( $n = 143$ ), and schools with more than 750 students ( $n = 96$ ). We used these size categories to describe the schools, teachers, and students in our samples. For multivariate analyses, we created two dummy-coded (0,1) indicators of middle-sized and large schools, each compared with small schools (the uncoded category).

We measure school demographic composition with three variables: low-income students (an average of 82% low-income students in this sample

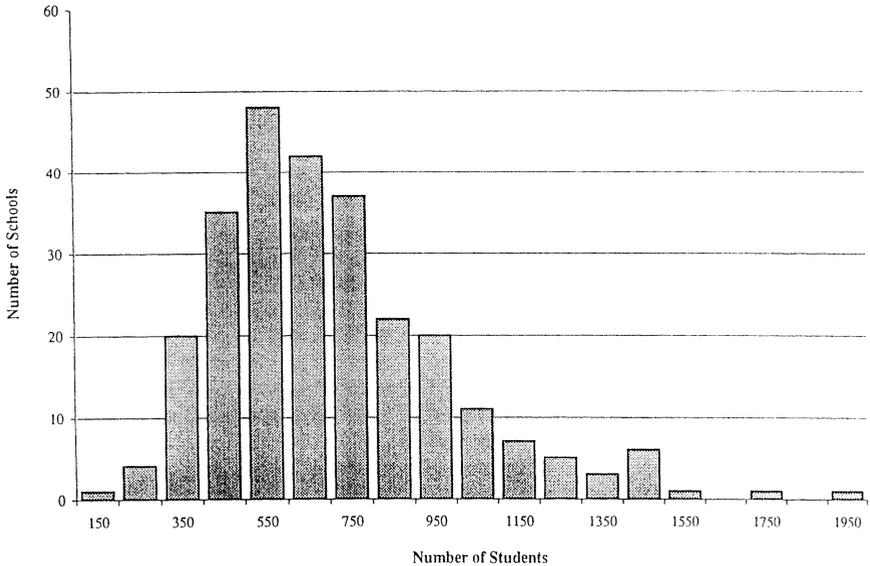


Figure 2. Histogram of the size distribution for elementary schools in the sample ( $n = 264$ )

of schools), Black students (54%), and Hispanic students (28%). Our multivariate models that focus on student achievement also include a measure of average collective responsibility, a school-level aggregate created from the teacher-level measure (see below) and then converted to a  $z$ -score ( $M = 0$ ,  $SD = 1$ ). More information about the construction of all variables included in our analysis is presented in the Appendix.

*Teachers.* The dependent measure in the multilevel analysis focusing on teachers and schools is collective responsibility. This composite was created with Rasch-equating methods. The measure taps the degree to which teachers believe their colleagues in the school share a responsibility for student learning and for students' academic and social development ( $M = 5.68$ ,  $SD = 2.17$ ). The measures of teacher background that we introduce for the purpose of statistical control include gender (77% female teachers), race/ethnicity (31% Black, 8% Hispanic, the remainder White or Asian teachers), and experience in the teaching profession (in years). The latter is represented with a series of dummy (0,1) variables: 1 year or less experience (3%); 2–3 years (8%); 4–5 years (8%); 6–10 years (14%); 11–15 years (10%); and the modal category, 16 or more years of experience (47%). We left the first category uncoded in multivariate analysis; thus, it is the group to which other experience categories are compared.

*Students.* The sample of students is approximately equally divided into sixth (51%) and eighth (49%) graders. The major dependent variable for students, their score on the 1997 mathematics portion of the Iowa Test of Basic Skills (ITBS), was administered in the winter of 1997.<sup>3</sup> The ITBS as-

esses general skills rather than knowledge of specific material. Alpha reliabilities of ITBS subtests, which measure the test's internal consistency based on a fall 1993 norming sample, ranges from .67 to .95. We were unable to find information about the construct or criterion validity of the ITBS. The mathematics portion of the ITBS tests students' understanding of the number system, mathematical terms and concepts, estimation, problem solving, and data interpretation. Salvia and Ysseldyke (1991) and <http://www.sde.state.id.us/instruct/SchoolAccount/PerformanceSummary/iowa.htm> provide additional information on the ITBS.

Students' annual test scores are part of the large database for the Chicago public schools available to researchers associated with the Consortium. The 1997 test scores were originally in the metric of grade-equivalents, but we wanted to combine the sample of sixth and eighth graders. Therefore, we transformed them into *z*-scores, standardizing each around its own grade-level mean (sixth or eighth). Before transforming, the grade-level equivalents were 6.6 for sixth graders and 8.6 for eighth graders. Our major control variable is students' scores on the same test the previous year (1996), also *z*-scored around grade-level means. Before transforming, these were 5.4 for sixth graders and 7.8 for eighth graders.

Besides the previous year's test scores, we include controls for students' SES (a *z*-score [ $M = 0$ ,  $SD = 1$ ]), dummy-coded race/ethnicity categories separately designating Black (the omitted category, 54% of this sample of students), Hispanic students (28%), and White and Asian students (17%). We also include dummy-coded (0,1) indicators of whether students were 2 or more years overage for their grade (16% were overage), a measure of mobility (i.e., the number of times the students had changed schools, which averaged .7 times), and the grade-level indicator.

## **Analytic Approach**

*Multilevel questions and methods.* The three research questions around which this study is organized are multilevel, consistent with our school effects approach. Addressing these questions involves estimating (a) the effects of school structure on teachers' attitudes (Question 1) and (b) the effects of school structure and social organization on student achievement (Questions 2 and 3). We used a multilevel analysis strategy, HLM (Bryk & Raudenbush, 1992). Its use in studies that investigate school effects is, by now, quite common. Below, we provide a brief description of the approach and describe its application to the three questions.

In general, HLM used in a school effects context involves three steps. In the first step, variance in the dependent measure is partitioned into its within- and between-school components. Only the proportion of variance in the outcome that lies systematically between schools may be modeled as a function of school characteristics. The second step estimates a within-school model for each school. Here, the dependent measure is estimated as a function of the characteristics of individual school members (in this case,

*Table 1*  
**Descriptive Statistics on Variables Measured on Schools, Teachers,  
and Students for the Chicago School Size Study**

	Small schools (<400 students)	Middle-sized schools (400–750 students)	Large schools (>750 students)
<b>A. Schools (<i>n</i> = 264)</b>			
Sample size	25	143	96
Percent low-income students	67.41 (23.24)	80.92 (20.69)	87.45 (12.03)
Percent black students	59.20 (38.23)	63.28 (40.83)	39.01 (41.71)
Percent Hispanic students	16.12 (22.20)	19.28 (27.41)	43.56 (35.56)
Percent White/Asian students	24.68 (25.18)	17.44 (24.93)	17.43 (23.23)
Average collective responsibility <sup>a</sup>	.643 (1.039)	.003 (1.050)	-.173 (0.813)
<b>B. Teachers (<i>n</i> = 4,495)</b>			
Sample size	260	2,066	2,179
Collective responsibility	5.68 (2.166)	5.72 (2.234)	5.70 (2.095)
Percent female teachers	76.15	76.80	77.61
Percent Black teachers	35.38	36.99	23.86
Percent Hispanic teachers	1.15	4.47	12.74
Percent White/Asian teachers	43.46	39.37	45.61
Teaching experience			
One year or less	2.69	3.11	3.03
2–3 years	7.31	7.86	8.23
4–5 years	5.38	8.59	8.64
6–10 years	12.69	13.69	14.34
11–15 years	9.23	10.39	10.67
16 years or more	52.69	47.86	45.84
<b>C. Students (<i>n</i> = 22,599, 6th and 8th graders)</b>			
Sample size	1,174	11,026	10,399
1997 mathematics test score <sup>b</sup>	.415 (.947)	.054 (1.032)	.012 (.972)
1996 mathematics test score <sup>b</sup>	.361 (.963)	.054 (1.033)	-.005 (.971)
Socioeconomic status <sup>a</sup>	.430 (.924)	.169 (.948)	-.008 (.929)
Percent female students	50.85	52.38	52.35
Percent Black students	64.22	65.06	41.13
Percent Hispanic students	14.31	17.38	42.15

*(Continued)*

*Effects of School Size on Attitudes and Achievement*

*Table 1 (Continued)*

	Small schools (<400 students)	Middle-sized schools (400-750 students)	Large schools (>750 students)
Percent White/Asian students	21.47	17.56	16.72
Percent Grade 8 (vs. Grade 6)	49.75	48.69	49.44
Percent overage <sup>c</sup>	13.37	16.18	16.29
Number of school changes	.606	.701	.655

*Note.* Standard deviation of continuous variables in parentheses.

<sup>a</sup>Variable is a z-score ( $M = 0$ ),  $SD = 1$ ) on the entire sample of schools, teachers, or students.

<sup>b</sup>Test scores were z-scored for 6th and 8th graders separately, around the means for each grade level. Z-scores were computed on a full sample of students and schools, rather than those with teacher data available.

<sup>c</sup>Overage students were born at least 2 years before the age cutoff for their grade.

teachers and students).<sup>4</sup> At the third step, the effect of school factors is estimated using between-school variance in the outcome, adjusted for characteristics of individuals in the school. The results of this third step provide estimates of the influence of school size on teachers' collective responsibility and students' mathematics achievement.

*Analysis among teachers.* The dependent measure, teachers' collective responsibility, was constructed as a Rasch-equated composite. This allows another step in the HLM model: a measurement model (Raudenbush, Rowan, & Kang, 1991), whereby collective responsibility is adjusted for measurement error in the composite. Thus, our analysis of teachers is a three-level HLM, with Level 1 being the measurement model. At Level 2 (within schools), we adjust collective responsibility for teacher demographics and experience. At Level 3 (between schools), we incorporate the social and demographic composition of the Chicago elementary schools.

*Analysis among students.* Students' scores on the 1997 ITBS mathematics score is the dependent measure in a two-level HLM. The intercept of this estimator, then, essentially serves as the dependent variable in a school-level model. At Level 1 (within schools), students' mathematics achievement scores are adjusted for their achievement on the same test the previous year (1996) as are demographic characteristics, mobility, and being overage. Thus, the dependent measure in the Level 2 analysis (between schools) represents a background-adjusted 1-year gain in mathematics for sixth and eighth graders. This modeling strategy investigates the effects of both school size and school social organization (measured by average collective responsibility) on mathematics learning, taking school social and demographic composition into account.

We are interested in whether school size influences learning directly and indirectly through teachers' attitudes (Question 3). Our final HLM model

on achievement estimates direct effects of size on students' mathematics learning. Moreover, if (a) school size has a direct effect on collective responsibility in the teacher-level analyses and (b) average collective responsibility is associated with mathematics learning in the student-level analyses, then we will have demonstrated that school size influences students. Our full HLM model on student achievement thus captures the idea of both a direct and an indirect effect.

*Presentation of results.* In order to compare size effects across two different outcomes in this study (and two different samples), we present the results in effect size (*SD*) units.<sup>5</sup> In this way, readers may compare the magnitudes of the effects in a common metric, beyond assessing their statistical significance. Our multilevel analyses focus on between-school differences; thus, we present results in between-school *SD* units.

## Results

### Descriptive Results for Small, Medium-Sized, and Large Schools

Distributions of all measures used in the study are displayed in Table 1. We present these descriptive statistics as subgroup means for small, medium, and large schools. We display means for schools, teachers, and students in these three school groupings. For continuous variables, we present both means and *SDs* (in parentheses below means); only group means are presented for dummy variables.

*Schools.* Schools enrolling less than 400 students represent only 9% of the sample (25 schools), whereas 36% of the schools enroll more than 750 students (96 schools; see Table 1, panel A). Fifty-four percent (143 schools) of Chicago elementary schools are middle-sized (400–750 students). Small schools differ from middle-sized and large schools in two major respects. First, they enroll fewer low-income students (although two-thirds of their students are from low-income families). Second, average collective responsibility is considerably higher than in the middle-sized and large schools. Proportions of low-income students differ between large and small schools by over 1 *SD* (a large effect). The difference in average collective responsibility between these two size groups is also large (.8 *SD*).

*Racial/ethnic compositions also differ by size.* The largest schools enroll fewer Black and more Hispanic students, whereas Black students are especially prevalent in middle-sized and small schools. We see that Chicago K-8 schools enroll mostly low-income (82%) students who are members of racial minority (75%) groups. However, students in small schools are characterized by relatively higher levels of social advantage, at least in comparison to other students in Chicago schools.

*Teachers.* Table 1 (panel B) shows that only 6% of the total sample of teachers work in small schools, whereas about equal proportions of the total sample of teachers work in middle-sized and large schools (46% and 48%, respectively). Among teachers, collective responsibility is higher in small than in medium and large schools (an advantage of about .25 *SD*). However,

### *Effects of School Size on Attitudes and Achievement*

the differences are not as large as the school average of this variable shown in Table 1, panel A. The proportion of Chicago elementary school teachers who are female (75%) is equally distributed across schools of different sizes. Chicago teachers are mostly non-minorities (31% Black; 8% Hispanic). The proportions of Black teachers are higher in small than in large schools, whereas the proportion of Hispanic teachers is highest in large schools (which is reasonable because large schools also enroll more Hispanic students; Table 1, panel C). Chicago's teachers are very experienced, but experience is unrelated to school size. One notable group difference is that the proportion of teachers with 16 or more years of teaching experience is lower in large (46%) than in small schools (53%); however, teachers in both groups are very experienced.

Teachers' demographic characteristics and experience differ only modestly by the size of the schools in which they teach. However, their attitudes about collective responsibility are more positive in small schools. It is possible that these group differences are related to the social composition of schools of different sizes.

*Students.* Mathematics achievement scores among students in the sixth and eighth grades differ considerably by school size (Table 1, panel C). On the 1997 mathematics test, students in small schools scored .40 *SD* above those in middle-sized and large schools, a moderate effect. Group differences, although slightly smaller for the 1996 test, follow a similar pattern: students in small schools scored .35 *SD* above those in medium and large schools. Of course, these achievement differences could be the result of relatively more advantaged students in small schools (students' SES is .3-.4 *SD* higher than their counterparts in larger schools). Student gender and grade level are balanced by size, but racial composition is not. Large schools enroll the lowest proportions of Black students (41%) and the highest proportion of Hispanic students (42%). Small schools, by contrast, enroll more Black (65%) and fewer Hispanic students (21%). Small schools enroll a higher proportion of White and Asian students (21%) than either middle-sized or large schools. In small schools, there are also lower proportions of overage students, and mobility is lower.

Chicago's small elementary schools seem considerably more advantaged than middle-size and large schools, in both academic and social terms. From the observed mean differences shown in Table 1, school size may be related to the outcomes for both teachers and for students. However, the advantage in terms of social composition that small Chicago elementary schools enjoy relative to their counterpart schools that enroll more than 400 students may well explain the observed differences in outcomes. Multivariate analyses should be helpful in sorting out these competing explanations.

*Making comparisons when using data from a population.* Most empirical studies use a sample of data from a population. As noted above, we are fortunate in being able to study a population. Under these circumstances, statistical tests of group differences may be redundant. Therefore, we have not reported statistical significance and probability levels for differences

*Table 2*  
**Fully Unconditional HLM Models: Psychometric Characteristics of  
 Dependent Variables in Chicago School Size Study**

---

A. Teachers: collective responsibility	
Within-school variance (sigma squared)	5.819 <sup>a</sup>
Between-school variance (tau)	.613
Between-school <i>SD</i>	.783
Reliability (lambda)	.731
Intraclass correlation <sup>b</sup>	.095
A. Students: 1997 mathematics achievement	
Within-school variance (sigma squared)	.774
Between-school variance (tau)	.237
Between-school <i>SD</i>	.485
Reliability (lambda)	.957
Intraclass correlation <sup>b</sup>	.233

---

<sup>a</sup>The results for collective responsibility are from a three-level hierarchical linear model (HLM). Level 1 includes adjustment for measurement error in the collective responsibility measure, constructed with Rasch-equating methods. Levels 2 and 3 are unconditional.

<sup>b</sup>The intraclass correlation (ICC) is the proportion of total variance in the outcome that lies systematically between schools. It is computed as follows:  $ICC = \tau / (\tau + \sigma^2)$ .

among school, teacher, and student subgroups in Table 1. However, we did not conduct such tests. Most group mean differences discussed above are statistically significant, particularly between small and large schools.

Although we know these differences exist for this population, we also note that our population is relatively small. Because of this, we also want to know whether the differences we observe for these Chicago students, teachers, and schools are representative of a larger population. Examples of the large populations to which we might generalize include students and teachers in Chicago elementary schools over all grade levels or over multiple time points, or perhaps to young adolescents in inner-city elementary schools elsewhere in the United States. Our population is surely not a perfect random sample drawn from these larger populations. Thus, we suggest that significance testing of the effects of independent variables provides useful information. We include tests of statistical significance for the multivariate results presented in the remainder of this article.

### Partitioning Variance in Teacher and Students Outcomes

Information from the fully unconditional HLM analyses of teachers and students is presented in Table 2. Only a modest proportion of the variance in teachers' attitudes about collective responsibility lies systematically between schools (intraclass correlation [ICC] = .095),<sup>6</sup> whereas a higher proportion of variance in students' 1997 mathematics achievement scores is found between schools (ICC = .233). In HLM, the reliability of the estimated school means for each outcome is estimated by the lambda statistic. Measured this

**Table 3**  
**Within-School HLM Model for Teachers' Collective Responsibility**  
**(*n* = 4,495 teachers in 264 schools)**

Independent variable	Beta coefficient
Intercept ( $\beta_{00}$ )	5.352**
Gender (female)	.094
Black <sup>a</sup>	.053
Hispanic <sup>a</sup>	-.003
Experience <sup>b</sup>	
2–3 years	-.136
4–5 years	-.378*
6–10 years	-.463**
11–15 years	-.377**
16 or more years	-.177

*Note.* In this hierarchical linear model (HLM), variance in the outcome varied between schools. Between-school variance in the control variables was fixed to zero. All control variables were centered around the grand mean. These decisions also apply to the results presented in Table 4.

<sup>a</sup>Black and Hispanic teachers were each compared with White/Asian teachers.

<sup>b</sup>Each experience group was compared with teachers with 1 year or less of experience.

\* $p \leq .01$ . \*\* $p \leq .001$ .

way, the reliability of teachers' collective responsibility is reasonable (.731), but more modest than that of student achievement, which is very high (.957). Differences in lambda reliability are, in part, a function of the average number of teachers per school (17) compared with the average number of students per school (86).

Both outcomes vary significantly between schools and both are reasonably reliable. Therefore, we proceed with the multilevel analyses to address our research questions. The psychometric properties of these two outcomes suggests, however, that our search for school effects is likely to be more successful among students than among teachers.

### **Multilevel Models of Teachers' Collective Responsibility**

*Within-school model.* Our Level-1 HLM model of teachers' collective responsibility factors out measurement error. Thus, the Level-2 HLM model explores effects of teachers' social background and experience on this outcome (Table 3). Neither teachers' gender nor their race/ethnicity is related to their assessments about collective responsibility.

However, teachers' experience is related to this outcome. Compared with teachers with a year or less of experience (the uncoded category), teachers with more experience have consistently less positive views about collective responsibility. Although a nonlinear trend, experience effects increase in magnitude (negatively) until the category of 6–10 years experience (Effect Size [ES] =  $-.59 SD$ ,  $p < .001$ ), level off for teachers with 11–15 years experience (ES =  $-.48 SD$ ,  $p < .001$ ), and decline somewhat for the most

Table 4  
**Full Between-School HLM Model for Teachers' Collective Responsibility (*n* = 4,495 teachers in 264 schools)**

Independent variable	Gamma coefficient
Intercept	6.947***
Percent low-income students	-.009*
Percent Black students	-.007*
Percent Hispanic students	-.003
School size	
Medium vs. small	-.406*
Large vs. small	-.589**

*Note.* To reduce the number of numbers presented, we chose not to present within-school coefficient for control variables in the hierarchical linear model (HLM; gender, race/ethnicity, and experience). They are almost unchanged from the results of the within-school model presented in Table 3.

\**p* ≤ .05. \*\**p* ≤ .01. \*\*\**p* ≤ .001.

experienced teachers (ES = -.23 *SD*, a nonsignificant difference). Although the effects of teaching experience on collective responsibility are not a focus of this study, the decline in this attitude as teachers gain more experience is striking. Moreover, the size of these effects is moderate to large.

*Between-school model.* The results of the full HLM model for collective responsibility are presented in Table 4. After controlling for school demographic composition, school size is significantly related to between-school

Table 5  
**Within-School HLM Model for Students' Achievement in Mathematics Responsibility (*n* = 22,599 students in 264 schools)**

Independent variable	Beta coefficient
Intercept ( $\beta_0$ )	.001
1996 mathematics achievement	.829**
Gender (female)	-.019*
White/Asian <sup>a</sup>	.093**
Hispanic <sup>a</sup>	.054**
SES	.038**
Grade 8 <sup>b</sup>	.031**
Overage	-.081**
Number of school changes	-.018**

*Note.* In this hierarchical linear model (HLM), variance in the outcome varied between schools. Between-school variance in the control variables was fixed to zero. All control variables were centered around the grand mean. These decisions also apply to the results presented in Table 6.

<sup>a</sup>White/Asian and Hispanic students were each compared with Black students.

<sup>b</sup>Students in Grade 8 compared with those in Grade 6.

\**p* ≤ .01. \*\**p* ≤ .001.

Table 6

**Full Between-School HLM Model for Students' 1997 Mathematics Achievement ( $n = 22,599$  students in 264 schools)**

Independent variable	Gamma coefficient
Intercept	.115*
Percent low-income students	-.002*
Percent Black students	.001
Percent Hispanic students	.011
Average collective responsibility	.026**
School size	
Medium vs. small	-.073*
Large vs. small	-.041

*Note.* To reduce the number of numbers presented, we chose not to present between-school parameters of this hierarchical linear model (HLM). They are almost unchanged from the results of the within-school model presented in Table 5.

\* $p \leq .05$ . \*\* $p \leq .01$ .

differences in teachers' attitudes about collective responsibility. Compared with small schools, teachers' attitudes in both medium-sized and large schools are significantly less positive. Teachers' views about the prevalence of collective responsibility in their schools are more negative in medium-sized than in small schools ( $ES = -.50 SD, p < .05$ ). Even more striking is the comparison between large and small schools ( $ES = -.73 SD, p < .01$ ).

Both size comparisons result in large and negative effects. Not only is school social composition taken into account in this model, but the social background and experience of the teachers are also accounted for. We did not find any differential effects for school size in schools with different social compositions.<sup>7</sup> Regarding Question 1, we conclude from this analysis that the size of Chicago elementary schools influences teachers' attitudes about collective responsibility and that small schools are favored.

**Multilevel Models of Students' Learning in Mathematics**

*Within-school model.* Results of the within-school model exploring student achievement are displayed in Table 5. The most important predictor of students' mathematics achievement scores in 1997 is their achievement status on the same test in the previous year ( $ES = 1.71 SD, p < .001$ ). The residual outcome is the 1-year gain in achievement between fifth and sixth grades or between seventh and eighth grades, interpreted as a measure of learning. Students' demographic characteristics are related to mathematics learning. Although the effect sizes are small in magnitude, all are significantly below the .01 probability level. Girls learn slightly less than boys ( $ES = -.04 SD$ ). SES is also positively related to learning ( $ES = .08 SD$ ). Compared with Black students (the comparison group), both Hispanic and White/Asian students learn more (.11 and .19  $SD$ , respectively).

Other features of students' background are also related to 1-year gains in mathematics scores, but again effect sizes are small. Eighth graders learn slightly more than sixth graders ( $ES = .06 SD$ ), students more than 2 years older than their classmates learn less ( $ES = -.17 SD$ ), and mobility has a small residual negative effect on learning ( $ES = -.04 SD$ ). This model indicates that even after taking previous achievement into account, Chicago middle-grade students' mathematics learning is associated with differences in their social backgrounds.

*Between-school model.* The full HLM model explores school effects on students' mathematics learning (Table 6). The model structure is similar to the model that explores teachers' collective responsibility in terms of the controls for school social composition. Besides estimating the effects of school size on learning, however, we also estimate the effect on learning of average collective responsibility. In response to Question 2, we find that this measure of school social organization is positively related to mathematics learning ( $ES = .19 SD, p < .01$ ), although the effect is small. Confirming our hypotheses, students learn more in schools where teachers have more positive attitudes about collective responsibility.

Question 3 is addressed in the same analysis, by considering the direct effects of school size on mathematics learning. Compared with small schools, students learn less mathematics in medium-sized schools ( $ES = -.54 SD, p < .05$ ), a large effect. Although the direction of the effect comparing small to large schools is also negative, the comparison is not statistically significant. We found that these size effects did not differ by school social composition.

The results of the full HLM model shown in Table 6 indicate that school size has both a direct and an indirect effect on student learning. The combined findings that (a) school size is strongly associated with teachers' attitudes about collective responsibility (Table 4), (b) average collective responsibility is related to student learning, and (c) school size is associated with learning (both from Table 6) lead us to conclude that size has an indirect effect on learning, through teachers' attitudes. The magnitude of the indirect effects is small, however:  $ES = -.10 SD$  for small versus medium-sized schools;  $ES = -.14 SD$  for small versus large schools.<sup>8</sup> The direct effects on learning are demonstrated by the school size effects in Table 6 ( $ES = -.54 SD$  for small vs. medium-sized schools;  $-.31 SD$  for small vs. large schools). This results in total effects favoring small schools on math learning of  $.64 SD$  compared with medium-sized schools,  $.45 SD$  compared with large schools. Results for these analyses of teachers and students in Chicago K-8 schools are consistent: smaller schools are favored.

## Discussion

### Why Would School Size Make a Difference?

*Small schools work better.* Both teachers and students are members of schools. Their schools are defined by many qualities and characteristics: who

goes there; who teaches there; how they are organized; the mission that underlies their operations; norms for how adults and students are expected to behave (including engagement, commitment, behavior, and productive activities); and how they are structured. This study focuses on a single structural characteristic, student enrollment. We examine whether this structural feature influences both teachers and students, after taking into account who goes there and who teaches there. This study demonstrates that school size really matters in Chicago. The findings are straightforward and consistent. Small elementary schools work better on the outcomes we examine. In schools with fewer than 400 students, teachers report that they and their colleagues assume more responsibility for student learning. Students attending smaller schools also learn more mathematics over a year.

*Size and social organization.* Although we designed our study to estimate direct effects of school size on these outcomes, it is difficult to see how the size of the school could actually influence students directly. Rather, size influences teachers and students indirectly. A reasonable mechanism might be that a smaller organizational dimension would facilitate personalized social interactions among school members. Teachers who interact more often with fewer students know their students better. By knowing students better, teachers are likely to worry more about their failures, provide more help directed toward improvement, take responsibility for disciplining everyone, and invest more fully in improving the whole school (these are some of the items included in the collective responsibility composite). This interpretation seems consistent with the indirect effects of school size on student achievement that we report.

Our study focuses on only one facet of school social organization, that is, collective responsibility. We find that this does influence learning. In a more complete analytic model, we believe that school size would have only an indirect effect. If we were to account for the many facets of school organization (particularly those defining their human dimensions) that are influenced by the number of students who attend, we suggest that such a model would completely “explain away” the direct effects of school size on learning. This surely does not mean that school size has no influence on learning; rather, in smaller schools, there are more intimate and personal social relations among teachers and students. The quality and character of these relationships are important determinants of student learning. This is a likely mechanism through which size affects learning.

### **Size Influences Teachers as Well as Students**

Most school effects studies focus on student outcomes, and this makes good sense. Virtually everyone would agree that the “major business” of schools is student learning. Teachers are a vital part of that business. The research base on school size, reviewed at the outset, is devoid of studies that explore how school size influences teachers. It could be argued that teacher outcomes like self-efficacy or job satisfaction are peripheral in a school effects context

because they have little influence on student achievement. However, it is harder to argue that teachers' attitudes about their students are unimportant, especially if those attitudes are shown to actually influence student learning.

In this study, we demonstrate that teachers' collective responsibility, as an organizational property of schools, has a positive influence on student learning. This association highlights the importance of exploring teacher outcomes in their own right. We suggest that research that estimates school effects on school members should focus on teachers as well as students. We are fortunate in this study to have access to data with information on large numbers of teachers and students in a substantial number of schools, a data structure that is, unfortunately, uncommon in educational research.

### **The Tradeoff Between Good Data and Low External Validity**

We acknowledge that the findings from this study have limited external validity. That is, we should not generalize our results beyond the public elementary schools in Chicago. To the extent that we can argue that Chicago's school population and school conditions reflect those of other large U.S. cities, perhaps the findings of this study are relevant for our nation's urban schools. Moreover, the organization of Chicago's schools is unusual (mostly elementary K–8 and 9–12 secondary schools, with very few stand-alone middle schools).

With these limitations in mind, some readers might question the utility of findings from a study of Chicago's K–8 schools to a research audience beyond the Chicago context (or even the urban school context). A large proportion of school effects studies, particularly those that use the appropriate analysis techniques, are secondary analyses (this study is no exception). Many of these studies have used data collected by the the U.S. Department of Education, most commonly studying high schools with NELS: 88 (National Education Longitudinal Study). All researchers who conduct secondary analyses make tradeoffs. Comparing analyses with the Chicago data to those using NELS:88 may be instructive.

The most obvious difference is generalizability. This study focuses on the public schools in one city, whereas NELS data are nationally representative (and include private schools). However, our data are not really samples, but include virtually the entire population. A related issue, one that also favors the Chicago data, is the large within-school sample sizes (for both students and teachers). Within-school student sample sizes in NELS are quite small (averaging less than 15 students per school). "Samples" of teachers are even smaller, and they were not selected separately from students. These limitations to teacher data from NELS render them inappropriate for school effects analysis (unless teachers are tied to students). With HLM, larger within-school sample sizes yield more precise and more stable estimates of the parameters under study. The large samples of students and teachers represent a substantial methodological strength of analyses using Chicago data.

Another advantage is level. In this study, we have expanded school effects studies by focusing on the elementary/middle grades, whereas it is only possible to study high schools with NELS.<sup>9</sup> Another comparative issue is time span: there are three waves of NELS data over the course of high schools, whereas the Chicago data focus on change over a single year. The structure of the Chicago data would, therefore, minimize school effects on change, because the period of change is quite short. However, the fact that we have identified significant effects of substantive magnitudes over such a short period suggests that these effects are real and would probably be considerably larger if the time period were longer. Given the high mobility rates among students in Chicago schools, it may not be realistic to focus on school effects over a period longer than a year without losing a sizable portion of the sample. We recognize that our analytic strategy, separating schools into categories by size, might also have some drawbacks. The group of 25 small schools that were favored in this analysis might share other advantages for which we have not accounted.<sup>10</sup>

Even after considering the several limitations of our analyses, we argue that the strengths of the database for Chicago schools outweigh the limitations. Moreover, most of the limitations would tend to produce underestimates of effects. Because a public-use version of Chicago survey data (as well as data from prior surveys) is available over the Internet, we suggest that researchers interested in urban schools have much to learn from analyzing this database. However, access to students' test scores is restricted to researchers associated with the Consortium for Chicago School Research.

### **How Do We Make Schools Smaller?**

Even though urban school populations are growing, it is unrealistic to think that there are many new (and smaller) schools on the drawingboards. Financial support through property taxes is a constant problem in large cities, although several states are tinkering with funding formulas to move toward a more equitable way to support education through the tax base. Although the policy implications of this study would recommend creating smaller elementary schools in the nation's inner cities, this must somehow be accomplished, in large part, within the reality of the structures that exist already.

A reform strategy that is becoming increasingly popular is to create smaller schools-within-schools within larger school buildings. There are many issues to be considered in implementing such a strategy. How small should the smaller units be? Although our categorization included schools with fewer than 400 students, we cannot say whether even smaller units might be more or less advantageous. What is the strategy for dividing children into units? We suggest that such designs should take great care not to increase between-unit stratification. Would dividing a large school into smaller units require more teachers, facilities, or funds (e.g., should the units share a music teacher, an art teacher, or a science lab)? How autonomous

should the separate units be? In schools with already small proportions of particular ethnic groups, should children be further divided? How would schools assure racial/ethnic and socioeconomic balance in the smaller units?

We return to an idea that was introduced earlier, that school size is particularly important in schools enrolling many disadvantaged students. That finding from the Lee and Smith (1997) study suggests that the issue of school size (at least in high schools) is especially important in urban schools. At the least, our findings suggest that school size is important for young adolescents' learning. They also suggest that size has an influence on teachers' attitudes about their students. Our findings provide empirical support for a move to small elementary schools.

## APPENDIX

### **Descriptive Information on All Variables Used in the Chicago School Size Study**

#### Measures Describing Schools

**SIZE:** We began with a variable describing the number of students in the school (Figure 2 shows its distribution). We created two dummy variables from the continuous measure: (1) size greater than or equal to 400 students but less than 750 students, and (2) size 750 students or more. The comparison group is schools enrolling fewer than 400 students.

**PERCENT LOW INCOME:** Percent of students in the school with low family income. "Low family income" is based on federal guidelines for poverty status.

**PERCENT BLACK:** Percent of African American students in the school.

**PERCENT HISPANIC:** Percent of Hispanic students (non-African American) in the school.

**AVERAGE COLLECTIVE RESPONSIBILITY:** Average of the Collective Responsibility measure across the teachers in each school (see below). Measure was z-scored ( $M = 0$ ,  $SD = 1$ ) for the 234 K-8 and PK-8 schools included in the Survey of Students and Teachers in the Chicago Public Schools.

#### Measures Describing Teachers

**COLLECTIVE RESPONSIBILITY:** This composite variable was computed as a Rasch-equated measure. It "focuses on the extent of a shared commitment among the faculty to improve the school so that all students learn. Teachers were asked how many of their colleagues feel responsible for students' academic and social development, set high standards of professional practice, and take responsibility for school improvement. A high score represents a strong sense of shared responsibility among the faculty who help each other reach high standards" (Bilcer, 1997, p.29). Forced response options for each item were "none,"

## *Effects of School Size on Attitudes and Achievement*

“some,” “about half,” “most,” and “nearly all.” The measure is composed of the following survey items, ranked in order of their “weight” in the composite:

How many teachers in this school:

- Feel responsible when students in the school fail?
- Feel responsible to help each other do their best?
- Help maintain discipline in the entire school, not just their classroom?
- Take responsibility for improving the school?
- Feel responsible for helping students develop self-control?
- Set high standards for themselves?
- Feel responsible that all students learn?

**BLACK:** Dummy variable coded 1 = African American, everyone else = 0. Residual comparison group composed of White or Asian teachers.

**HISPANIC:** Dummy variables coded 1 = non-African American Hispanic, everyone else = 0. Comparison group composed of White or Asian teachers.

**EXPERIENCE:** A series of five dummy variables for the number of years the teacher has been teaching: (1) 2–3 years, (2) 4–5 years, (3) 6–10 years, (4) 11–15 years, (5) 16 or more years. The comparison group are those who have taught one or fewer years.

### Measures Describing Students

**1997 MATHEMATICS SCORE:** Scores for 6th and 8th graders combined into a single scale. Each student's grade equivalent score on the mathematics section of the Iowa Test of Basic Skills (ITBS) administered to all Chicago elementary/school students each year. Score converted to a z-score ( $M = 0$ ,  $SD = 1$ ), centered around 6th or 8th grade mean.

**1996 MATHEMATICS SCORE:** Scores for 6th and 8th graders combined into a single scale. Student's Rasch-equated mathematics score from the 1996 ITBS, scaled for item difficulty, converted to a z-score ( $M = 0$ ,  $SD = 1$ ), centered around each grade-level mean.

**GRADE 8:** Dummy variable coded 1 = student is an 8th grader, 0 = student is a 6th grader. Approximately one half of the sample is 6th graders.

**SES:** A composite measuring family socioeconomic status (z-scored  $M = 0$ ,  $SD = 1$ ). Composite include parents' education (higher of mother's or father's, if student lives with both parents) and the number of educationally related possessions in the home. Unfortunately, students' family income (typically a component of SES) was unavailable in the dataset.

**WHITE/ASIAN:** Dummy variable coded 1 = student is White or Asian, 0 = everyone else. Comparison group is African American students.

**HISPANIC:** Dummy variable coded 1 = student is non-African American, Hispanic, or

Native American, 0 = everyone else. Comparison group is African American students.

MOBILITY: The number of times the student has changed schools, since s/he entered the Chicago Public Schools.

OVERAGE: Dummy variable coded 1 if the student is 2 or more years older than s/he should be in his/her grade, 0 otherwise. For 6th graders, the cutoff was September 1986; for 8th graders the cutoff was January 1984.

## Notes

Support for this study came from the Spencer Foundation, which we gratefully acknowledge. We appreciate the access to data from the Consortium on Chicago School Research (CCSR) and its Senior Director, Anthony S. Bryk. We received helpful advice on using the data from Diane King Bilcer and Julia B. Smith; however, any errors in using the data or interpreting the results are ours alone. The surveys were constructed in coordination with the Chicago Annenberg Study Project, which is associated with CCSR. The first author is a member of the Chicago Annenberg Study Project research team. For more information, contact [velee@umich.edu](mailto:velee@umich.edu)

<sup>1</sup>It may seem that summer learning loss is relevant here. Cooper, Nye, Charlton, Lindsay, and Greenhouse (1996), who conducted a research synthesis of 39 studies, concluded that students' test scores average at least 1 month lower when they return to school in the fall than when they left in the spring. Entwisle and Alexander (1992) present convincing evidence that the summer learning loss is greater for minority and low SES children in the early grades. However, they also found that students who lose the most between spring and fall also gain the most during the first few months of school, when they relearn much of what they lost over the summer. Were we to investigate test score differences between fall and spring, we would surely overestimate the yearly gain of the students who lose the most over the summer. If schools differed in their influence on summer learning loss, this effect would be problematic in our study. However, our analytic models control for test scores from the previous year, which were measured during the same month (April). We suggest, therefore, that our models provide a good estimate of 1-year gains in achievement. As Chicago does not test its students in both fall and spring, we are unable to investigate summer learning loss in this study.

<sup>2</sup>Ideally, in a school effects study with data from students, teachers, and schools, we would pose a three-level nested model, with students nested in classrooms and classrooms nested in schools. Unfortunately, in the 1997 Chicago survey data, teacher identifications are not linked to the students they taught. Rather, both are linked through the schools to which they belonged.

<sup>3</sup>The ITBS, as administered to Chicago students, tests mathematics and reading comprehension. We ran our full HLM analytic models for both subjects. Because the results were quite similar across both subjects, we chose to present results only for mathematics. Results for the analysis on reading are available from the authors on request.

<sup>4</sup>In this analysis, we did not model any distributional outcomes (i.e., we did not estimate any slopes as outcomes). As a result, the independent variables measuring individual characteristics were included as statistical controls. Thus, we constrained the between-school variance in these relationships to zero (i.e., they were "fixed" in the HLM models). We centered these variables around their respective grand means.

<sup>5</sup>Presentation and discussion of empirical results in effect-size units are common. We follow the standards suggested by Rosenthal and Rosnow (1984, p.360): effects that are .5 SD or more in magnitude are considered large; effects in the .3-.5 SD range are moderate; effects that are .1-.3 SD are small; and those below .1 SD are trivial. We find it useful to compare effect size magnitudes using these guidelines, which allows us to make substantive interpretations from group mean differences.

<sup>6</sup>Although the individual survey items that comprise the collective responsibility com-

## *Effects of School Size on Attitudes and Achievement*

posite measure queried teachers' views about their colleagues (hence, "collective" responsibility), the fact that so little of the variance in this outcome lies between schools (9.5%) suggests that the measure is probably tapping teachers' own views about these issues. In the study, our analyses of teachers suggest that we have interpreted this measure as a measure of each teacher's attitudes about his or her students. However, in the analyses of student achievement, our use of the aggregate suggests the "collective" interpretation.

<sup>7</sup>The Lee and Smith (1997) study of high school size reported that school size had stronger effects on achievement in schools enrolling students who were more disadvantaged (i.e., in lower SES and high-minority enrollment schools). Differential size effects were computed with interaction terms in that study. Their findings led us to compute and test interactions between the two school-size dummy variables and our school composition measures (percent low income, percent Black, and percent Hispanic students) on both collective responsibility and mathematics achievement scores. In neither analysis were these interaction terms statistically significant, either as a set or individually.

<sup>8</sup>We computed the indirect effects in the model shown in Table 6 as the product of the direct effects of school size on mathematics achievement (-.544, -.305) and the direct effect of average collective responsibility on learning in the same model (.194 *SD*).

<sup>9</sup>The NELS base-year study is cross-sectional; data were on students as eighth graders, when they were in middle-grade schools. However, the NELS base-year data contain no measures that allow researchers to control for prior ability. Thus, their appropriateness for school effects studies is limited.

<sup>10</sup>Given that the group of small schools is not numerous, and that they are all located in a single city, it might be fruitful to conduct field research in these schools to further investigate the mechanisms through which size translates into learning. This would be impossible in a national school sample, for both confidentiality and logistical reasons.

## References

- Barker, R., & Gump, R. (1964). *Big school, small school: High school size and student behavior*. Stanford, CA: Stanford University Press.
- Bilcer, D. K. (1997). *User's manual, version 1. Improving Chicago's schools: The teachers turn, 1997; The students speak, 1997*. Chicago: Consortium on Chicago School Research, University of Chicago.
- Brattesani, K. A., Weinstein, R. S., & Marshall, H. H. (1984). Student perceptions of differential treatment as moderators of teacher expectation effects. *Journal of Educational Psychology, 76*(2), 236-247.
- Brophy, J. E. (1983). Research on the self-fulfilling prophesy and teacher expectations. *Journal of Educational Psychology, 75*, 631-661.
- Brophy, J. E. (1986). *Socializing students' motivation to learn*. East Lansing: Michigan State University Press.
- Bryk, A. S., & Driscoll, M. E. (1988). *The school as community: Theoretical foundations, contextual influences, and consequences for students and teachers*. Madison: Center on Effective Secondary Schools, University of Wisconsin.
- Bryk, A. S., Lee, V. E., & Holland, P. B. (1993). *Catholic schools and the common good*. Cambridge: Harvard University Press.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.
- Bryk, A. S., Sebring, P. B., Kerbow, D., Rollow, S., & Easton, J. Q. (1998). *Charting Chicago school reform: Democratic localism as a lever for change*. Boulder, CO: Westview Press.
- Buzacott, J. A. (1982). *Scale in production systems*. New York: Pergamon.
- Carnegie Task Force on Teaching as a Profession. (1986). *A nation prepared: Teachers for the 21st century*. New York: Carnegie Forum on Education and the Economy.

## Lee and Loeb

- Chambers, J. G. (1981). An analysis of school size under a voucher system. *Educational Evaluation and Policy Analysis*, 3, 29–40.
- Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greenhouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, 99, 227–268.
- Cooper, H. M., Findley, M., & Good, T. (1982). Relations between student achievement and various indexes of teacher expectations. *Journal of Educational Psychology*, 74(4), 577–589.
- Cooper, H. M., & Tom, D. Y. (1984). Teacher expectation research: A review with implications for classroom instruction. *Elementary School Journal*, 85(1), 77–89.
- Entwisle, D. R., & Alexander, K. L. (1992). Summer setback: Race, poverty, school composition, and mathematics achievement in the first two years of school. *American Sociological Review*, 57, 72–84.
- Etzioni, A. (1975). *A comparative analysis of complex organizations*. New York: Free Press.
- Firestone, W. A., & Herriot, R. E. (1982, December). Prescriptions for effective elementary schools don't fit secondary schools. *Educational Leadership*, 40(3), 51–53.
- Firestone, W. A., & Rosenblum, S. (1988). Building commitment in urban schools. *Educational Evaluation and Policy Analysis*, 10, 285–300.
- Firestone, W. A., & Wilson, B. L. (1985). Using bureaucratic and cultural linkages to improve instruction: The principal's contribution. *Educational Administration Quarterly*, 21(2), 7–30.
- Gamoran, A. (1989). Measuring curriculum differentiation. *American Journal of Education*, 97(2), 129–143.
- Garbarino, J. (1980). Some thoughts on school size and its effects on adolescent development. *Journal of Youth and Adolescence*, 9(1), 19–31.
- Greene, C. N., & Podsakoff, P. M. (1981). Effects of withdrawal of performance contingent reward of supervisory influence and power. *Academy of Management Journal*, 24, 527–542.
- Guthrie, J. (1979). Organizational scale and school success. *Educational Evaluation and Policy Analysis*, 1(1), 17–27.
- Johnson, S. M. (1990). Redesigning teachers' work. In R.F. Elmore and Associates (Eds.), *Restructuring schools: The next generation of educational reform* (pp. 125–151). San Francisco: Jossey-Bass.
- Kenny, L. (1982). Economies of scale in schooling. *Economics of Education Review*, 2(1), 1–24.
- Lee, V. E., & Bryk, A. S. (1988). Curriculum tracking as mediating the social distribution of high school achievement. *Sociology of Education*, 61(2), 78–94.
- Lee, V. E., & Bryk, A. S. (1989). A multilevel model of the social distribution of high school achievement. *Sociology of Education*, 62, 172–192.
- Lee, V. E., Bryk, A. S., & Smith, J. B. (1993). The organization of effective high schools. In L. Darling-Hammond (Ed.), *Review of Research in Education*, 19, (pp. 171–267) Washington, DC: American Educational Research Association.
- Lee, V. E., & Smith, J. B. (1996). Collective responsibility for learning and its effects on gains in achievement for early secondary school students. *American Journal of Education*, 104, 103–147.
- Lee, V. E., & Smith, J. B. (1997). High school size: Which works best, and for whom? *Educational Evaluation and Policy Analysis*, 19(3), 205–227.
- Lieberman, A. (Ed.). (1992). *The changing context of teaching. Ninety-first yearbook of the National Society for the Study of Education* (Vol. 1). Chicago: University of Chicago Press.

### *Effects of School Size on Attitudes and Achievement*

- Little, J. W. (1982). Norms of collegiality and experimentation: Workplace conditions of school success. *American Educational Research Journal*, 19, 325–340.
- Little, J. W. (1990). The persistence of privacy: Autonomy and initiative in teachers' professional relations. *Teachers College Record*, 91(4), 509–536.
- Lortie, D. C. (1975). *Schoolteacher: A sociological study*. Chicago: University of Chicago Press.
- Louis, K. S., & Kruse, S. D. (1995). *Professionalism and community: Perspectives on reforming urban schools*. Thousand Oaks, CA: Corbin.
- Louis, K. S., Marks, H. M., & Kruse, S. D. (1996). Teachers' professional community in restructuring schools. *American Educational Research Journal*, 33(4), 757–798.
- Michelson, S. (1972). Equal school resource allocation. *Journal of Human Resources*, 7, 283–306.
- Monk, D. (1987). Secondary school size and curriculum comprehensiveness. *Economics of Education Review*, 6, 137–150.
- Monk, D., & Haller, E. J. (1993). Predictors of high school academic course offerings: The role of school size. *American Educational Research Journal*, 30, 3–21.
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: Government Printing Office.
- Oakes, J. (1985). *Keeping track: How schools structure inequality*. New Haven: Yale University Press.
- Raudenbush, S. W. (1984). Magnitude of teacher expectancy effects on pupil IQ as a function of the credibility of expectancy induction: A synthesis of findings from 18 experiments. *Journal of Educational Psychology*, 76(1), 85–97.
- Raudenbush, S. W., Rowan, B., & Kang, J. S. (1991). A multilevel, multivariate model for studying school climate with estimation via the EM algorithm and application to U.S. high school data. *Journal of Education Statistics*, 6, 295–330.
- Rosenholtz, S. J. (1989). *Teachers' workplace: The social organization of schools*. New York: Longman.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom*. New York: Holt, Rinehart, and Winston.
- Rosenthal, R., & Rosnow, R. L. (1984). *Essentials of behavioral research: Methods and data analysis*. New York: McGraw-Hill.
- Rowan, B. (1990a). Applying conceptions of teaching to organizational reform. In R. F. Elmore & Associates (Eds.), *Restructuring schools: The next generation of educational reform* (pp. 31–58). San Francisco: Jossey-Bass.
- Rowan, R. (1990b). Commitment and control: Alternative strategies for the organizational design of high schools. In C. B. Cazden (Ed.), *Review of research in education* 16, (pp. 353–389). Washington, DC: American Educational Research Association.
- Salvia, J., & Ysseldyke, J. E. (1991). *Assessment* (5th ed.). Boston: Houghton-Mifflin.
- Weber, M. (1947). *Theory of social and economic organization* (A. M. Henderson & T. Parsons, Trans.). New York: MacMillan.
- Wyner, N. B. (1991). *Current perspectives on the culture of schools*. Cambridge, MA: Brookline.

Manuscript received September 1, 1998

Revision received April 27, 1999

Accepted June 29, 1999