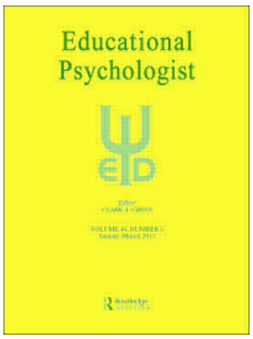

Educational Psychologist

ISSN: 0046-1520 (Print) 1532-6985 (Online) Journal homepage: www.tandfonline.com/journals/hedp20

Using Hierarchical Linear Modeling to Study Social Contexts: The Case of School Effects

Valerie E. Lee

To cite this article: Valerie E. Lee (2000) Using Hierarchical Linear Modeling to Study Social Contexts: The Case of School Effects, Educational Psychologist, 35:2, 125-141, DOI: 10.1207/S15326985EP3502_6

To link to this article: https://doi.org/10.1207/S15326985EP3502_6

EDUCATIONAL PSYCHOLOGIST, *35*(2), 125–141

Using Hierarchical Linear Modeling to Study Social Contexts: The Case of School Effects

Valerie E. Lee
School of Education
University of Michigan

This article focuses on a particular educational context, the school, and how characteristics of the structure and organization of high schools influence students' academic development. The emphasis is on a type of quantitative inquiry called *school effects research.* It describes a methodology that is most appropriate for conducting studies of school effects in particular and educational contexts in general: hierarchical linear modeling (HLM). Two previously published studies are used as heuristic examples of school effects studies conducted with HLM methods. Both studies use large and nationally representative longitudinal data from the National Education Longitudinal Study of 1988 to explore school effects on learning and its social distribution by student socioeconomic status. Study 1 focuses on the effects of high school size on learning. Study 2 focuses on how teachers' attitudes, taken as a collective property of the social organization of schools, influence both learning and its social distribution. Implications for both policy and research are discussed.

As the authors of the articles in this special issue argue collectively, children's learning is strongly influenced by the educational context in which it occurs. When children are very young, the family is the major context for learning. As they grow, children's educational activities occur increasingly in more formal settings. Within the context of preschool and elementary school, young children experience much of their education in small groups, either grouped into classrooms or into smaller groups within classrooms (often differentiated by skill level in reading or mathematics). Each grouping—indeed, almost any setting where formal or informal learning experiences occur—represents an educational context. Self-selected peer groups—groupings organized around extracurricular activities, sports teams, clubs, band, chorus—represent contexts in which formal and informal learning occur.

The nature of formal educational contexts also change as children mature. Younger children spend much of their school time in a single classroom, with a modest number of peers and a single teacher who instructs them in several subjects. As adolescents, children still experience education in classrooms, but not the same class all day. In adolescents' classrooms, students are exposed to several teachers and a more numerous and diverse group of classmates. The logic of specialization drives such changes in the organization of schooling. As the knowledge to be imparted is more complex as students mature, older children are best served by teachers with specialized expertise in particular subjects. Most high school students visit several classrooms every day, each staffed by a different teacher.

In two different ways, this article attempts to simplify what could be a complex discussion. Compared to that of elementary school children, adolescents' learning contexts are varied and complex. Thus, one simplification is to focus on the school itself as an organizational unit able to define the major educational context for adolescents. Although schools may influence students in many ways, a second simplification is to focus on academic outcomes. More precisely, I concentrate on learning, represented as a change over time in high school students' achievement. Therefore, the focus is on school-by-school differences in how much students learn. It is a well-established phenomenon, however, that how much students learn over time is influenced by their own background characteristics (e.g., ability, race, social class, gender). I introduce some complexity to the discussion by expanding the definition of academic outcomes. Besides investigating learning as a function of the characteristics of the schools students attend, I also explore how schools influence the social distribution of achievement. As Lee and Bryk

Requests for reprints should be sent to Valerie E. Lee, School of Education, University of Michigan, Ann Arbor, MI 48109. E-mail: velee@umich.edu

(1989) stated, "A compelling educational puzzle involves discovering why some schools are better able to induce academic outcomes among a broad social and racial distribution of students than others" (p. 172). Exploring learning and its social distribution as functions of school characteristics combines the common notions of both excellence (i.e., high achievement or learning) and equity (i.e., learning that is relatively unrelated to students' social background).

Rather than presenting new empirical research in this article, my aim is to explore the school context within both the conceptual and the methodological issues involved in the investigation of school effects. This line of inquiry, now quite common among sociologists of education, is somewhat less familiar to educational psychologists. What may be new to this readership is the methodology employed to investigate contextual effects. Among sociologists of education, school effects studies using quantitative methods typically employ hierarchical linear modeling (HLM) methods (Bryk & Raudenbush, 1992). I argue that this approach is both the most appropriate and also the most useful for the study of contextual effects, particularly the context of school. Because I have conducted many studies that fit the school effects model, virtually all of which make use of the HLM methodology, I use results from two of my published studies to illustrate how high school contexts may influence students' learning. I also use this venue to provide some explanation of the HLM methodology for those who are not familiar with it.

BACKGROUND

Three Lines of Relevant Research

Cross-sector comparisons. The search for school effects is embedded in three streams of theory and research, each developed independently of the others. My work on school effects began with explorations of the relative effectiveness of Catholic and public high schools. Much controversy arose around the findings of James Coleman and his colleagues (e.g., Coleman & Hoffer, 1987; Coleman, Hoffer, & Kilgore, 1982; Greeley, 1982; Hoffer, Greeley, & Coleman, 1985) who argued that students in Catholic high schools outperformed their public school counterparts. Moreover, these same researchers reported that the relation between social background and academic achievement was weaker in Catholic than public schools.

Evidence from both field and quantitative research in Catholic high schools (Bryk, Lee, & Holland, 1993; Lee & Bryk, 1988; Lesko, 1988) suggested that the academic organization and normative environments of these schools distinguish them from their public school counterparts. Thus, cross-sector comparative research began to suggest that organizational characteristics of schools influence both students' performance within schools and how that performance is distributed across students with different social backgrounds.

Effective schools. A line of inquiry often referred to as *effective schools research* is also relevant. These studies attempted to identify the characteristics of schools that make them instructionally effective for disadvantaged students (e.g., Brookover, Beady, Flook, Schweitzer, & Wisenbaker, 1979; Clark, Lotto, & McCarthy, 1980; Edmonds, 1979; Purkey & Smith, 1983; Rosenholtz, 1985; Rutter, Maughan, Mortimore, Outson, & Smith, 1979). This research, often targeted more to practitioners than researchers, was typically formulated in a two-stage process. The first stage identified schools that are particularly effective for low-socioeconomic status (SES) children. In the second stage, researchers searched for characteristics that were common among the schools identified as effective.

Although this line of research received much criticism on both methodological and substantive grounds and not all studies agreed on the salient factors for the effective schooling of disadvantaged children, the studies identify several factors as important. Effective schools have strong leadership focused on academic outcomes. They closely monitor student work. In such schools, teachers hold positive expectations for all students. Their social environments are purposeful. Their climates are orderly. Even though findings from the effective schools research have intuitive appeal, evidence for the claimed schools effects has not been solid.

Bureaucratic versus communal organization. The two streams described previously may be embedded in a theoretical perspective, much broader than the others, that is primarily sociological. This perspective contrasts two models of schooling: the rational–bureaucratic and the personal–communal. Eloquently articulated in an educational context by Bidwell (1965), the models are based on a fundamental sociological distinction set out by Weber (1926/1947), expanded on by Merton (1949), Parsons (1951), and Sorokin (1928), and applied to schools by Waller (1932). From the bureaucratic perspective, schools (especially high schools) are formal organizations, characterized by a functional division of adult labor into specialized tasks. Teaching roles are defined by subject matter and types of students. Social interactions are rule driven and affectively neutral. Authority is attached to roles rather than to the individuals who occupy the roles.

By contrast, the communitarian perspective views schools as small societies, organizations emphasizing informal and enduring social relationships, and driven by a common ethos. Two consequences of a communal organization are that the role of adults is typically quite diffuse and the division of labor minimal. This theoretical perspective was used by Lee, Bryk, and Smith (1993) to organize their extensive review. I suggest that a contrast of the specialization and communal models is a useful organizational lens through which to examine contextual effects in secondary schools.

More Specific Background

The theoretical basis for school effects studies lies in one or often several of these research perspectives. Of course, the types of school effects explored in individual studies (e.g., academic organization, social organization, school structure, social composition) also have their own literatures. Moreover, each outcome (e.g., academic achievement, learning, educational aspirations, attitudes toward school, engagement in school) is also embedded in broad research streams. The purpose of this very brief review of the research perspectives is to introduce ideas that are probably less familiar to readers whose primary interest is in educational psychology.

ORGANIZATION OF THIS ARTICLE

The remainder of this article is divided into four sections. First, I provide a brief explanation of the HLM methodology, particularly in terms of exploring the social context of schools. My purpose is less to provide statistical details of the technique than to explain the advantages of HLM over conventional statistical methods when investigating multilevel questions and to provide readers with enough information to understand the examples that follow. Readers interested in more detail about the statistical properties of HLM are referred to other sources. The second and third sections describe two studies that use HLM to explore school context effects. Again, the presentation is meant to be heuristic rather than either technical or explanatory. Although descriptive information about the samples used for each study is presented in numerical form, multivariate and multilevel results of both studies are presented in graphic form. The graphs focus only on the major research questions (which examine how school contexts influence student learning) without presenting effects of confounding variables. Drawn from published research, these are examples of school effects studies using the HLM methodology. Readers interested in more detail about the studies are referred to the articles from which they were drawn. Finally, I revisit the issue of contextual effects, again through the school effects lens, to argue for the usefulness (indeed the necessity) of this approach to explore how educational contexts influence students.

OVERVIEW OF THE HLM METHODOLOGY

HLM Solves Several Difficulties With Previous Research

Units of analysis. Empirical investigations of school effects, by their very nature, are hierarchical because students are nested in schools (i.e., several individuals experience their educations in the same school). These studies involve a search for statistical associations between school factors, on one hand, and variables measured on students on the other. Prior to the availability of multilevel methods in general and HLM in particular, there was considerable conceptual and methodological difficulty in conducting this type of research. A major issue involves answering the standard question: What is the appropriate unit of analysis for such studies? Prior to HLM, there were two typical approaches, neither of which was appropriate. Data were analyzed either at the level where the intervention or effect was administered (i.e., the school level) or at the level where the effect was thought to occur (i.e., the student level).

An example may be useful here. Let me suggest that a researcher was interested in investigating the influence of the racial composition of the school (which is quite separate from a student's race or ethnicity) on students' academic achievement. If the researcher were to ignore the essential multilevel nature of the question, he or she would analyze the data using conventional single-level methods, such as ordinary least squares (OLS) regression or analysis of variance (ANOVA). Typically, the researcher may simply append a variable measuring school racial composition onto a file that also includes measures of students' race, SES, and ability. If school racial composition were measured on an interval scale, the researcher would select OLS regression; if the measure were in categories, ANOVA would be the technique of choice.

Were the researcher to decide to analyze the data at the individual (i.e., student) level, he or she would simply append a measure of school racial composition onto each student in a particular school. This technique would assume that the achievements of all students in the school were influenced identically by school racial composition. Another approach would be to analyze the data at the school level. Here the analyst would use the same methods but would simply aggregate the data to the school level. This approach would ignore the substantial variability that exists in the dependent variable as well as the independent variables. I argue that neither approach is appropriate; both the student and the school must be considered as important units of analysis. HLM allows researchers to consider more than one unit of analysis, using multilevel data.[1]

Three problems, one solution. Three major difficulties commonly occur in the analysis of multilevel data and questions with single-level methods such as OLS regression, ANOVA, or structural equation modeling: aggregation bias, misestimated standard errors, and heterogeneity of regression. Aggregation bias can occur when a variable takes on different meanings and, therefore, has different effects at different levels of aggregation. For example, the average SES of a school may influence a student's academic achievement above and beyond his or her own SES. In the

[1]Readers who wish more detail about the HLM methodology are referred to Bryk and Raudenbush (1992), a text that provides complete explanation. The brief explanation of HLM provided in this article draws heavily from Lee and Bryk (1989). The purpose of that article, one of the first published studies using HLM for a substantive analysis, was also meant to provide heuristic models for the analyses guiding that study.

example given previously, school racial composition is clearly different than an individual student's race or ethnicity, and the effects of the two variables on student achievement may be quite different.

A second difficulty concerns the estimation of the standard errors used for statistical testing. With multilevel data, misestimated standard errors can occur when researchers treat individual cases as though they are independent (a standard assumption of OLS regression methods) when they are not. For example, students' achievement in the same school shares at least some dependence with other schoolmates. This dependence arises from two phenomena: Students in the same school share many experiences, and students are assigned to schools in a nonrandom manner, typically by residential location. In a single-level analysis, we must assume that relations among students' personal characteristics (e.g., race, ethnicity, or SES) and outcomes (e.g., achievement) are similar, regardless of the schools students attend. In general, the standard errors of the school-level variables in an analysis such as the one described previously would be overestimated in a single-level analysis at the individual level, which would result in an underestimate of the statistical significance of the group-level effects. It has been the case that school effects, when estimated in an individual level analysis, have been consistently underestimated.

A third difficulty concerns heterogeneity of regression slopes. That is, relations between characteristics of students (such as race, ethnicity, SES) and academic achievement may vary across schools and may be functions of group-level variables. For example, the achievement gap between students of different racial groups could vary considerably between schools, and such a regression coefficient could depend on the racial composition of the school. Research prior to the development of HLM has typically looked on this phenomenon as a methodological nuisance rather than an interesting phenomenon. The causes of such heterogeneity may, however, actually be of substantive interest within the school-effects context. For example, the relation between SES and achievement, which can be viewed as a measure of social inequity in schools, may be associated with particular organizational features of schools.

Compared to conventional single-unit methodologies such as OLS regression or ANOVA, the use of HLM for analysis of multilevel questions, with multilevel data, solves each of these difficulties. Aggregation bias is avoided because we are able to investigate the effects of a similar phenomenon (e.g., SES) at more than one level of aggregation (e.g., school average SES and students' SES). The assumption of independence of cases is not necessary in a multilevel analysis, in that the probable dependence of students who experience their educations in the same school is explored explicitly with nested data. We may also use HLM to explore whether relations between social background and outcomes (e.g., the social distribution of achievement by student SES) vary between schools. If they do, we may investigate school factors that are associated with the social distribution of achievement within schools.

Conducting an HLM Analysis

Three steps. There are three steps in a typical HLM model used in a school-effects study. In the first step, the researcher partitions the variance in a dependent variable, which is measured on individual students, into two parts—the proportion of variance that lies between students in the same school (pooled over schools) and the proportion of variance in the dependent variable that lies systematically between schools. Only the proportion of variance in the dependent variable that occurs between schools may be modeled as a function of school characteristics. Determining the proportion of the total variance that lies systematically between schools, called the intraclass correlation (ICC), constitutes the first step in an HLM analysis.[2] We conduct this analysis with a fully unconditional model, which means that no student or school characteristics are considered. This first step can also indicate whether HLM is needed or whether a single level analytic method is appropriate. Only when the ICC is more than trivial (i.e., greater than 10% of the total variance in the outcome) would the analyst need to consider multilevel methods. Ignoring this step (i.e., assuming an ICC of either 0 or 1) would be inappropriate if the research question were multilevel. Investigation of contextual effects, I argue, is by nature a multilevel question.

The second step in an HLM school-effects analysis involves the estimation of a within-school or Level 1 model. In this step, we investigate the characteristics of individual students that are associated with the outcome. Typical outcomes in a school-effects study at the secondary level include achievement or learning, the number and types of courses taken in particular subjects, dropping out prior to graduation, educational aspirations, academic self-concept, or engagement with school. Student characteristics could be divided into demographics (e.g., race, ethnicity, gender, SES), ability (e.g., an IQ measure or prior achievement), attitudes (e.g., motivation or interest in school), and behaviors (e.g., attendance, disciplinary problems, time spent on homework, participation in extracurricular activities).

Continuing with the achievement example, we could explore this outcome as a function of such student characteristics as ability or prior achievement, race, ethnicity, SES, gender, the courses students have taken, or their expectations

[2] In a single-level analysis at the student level, we would assume that the ICC would be 0. In a single-level analysis at the school level, we would assume that the ICC would be 1. Obviously, neither is correct. Typical multilevel analyses using achievement as an outcome have found the ICC to be about 25% to 30% (i.e., about 25% to 30% of the total variance in achievement lies systematically between schools).

for future education. Moreover, we may consider these variables measured on students as either statistical controls or social distribution parameters that we investigate as functions of school characteristics (e.g., the relation between SES and achievement). Within a Level 1 HLM analysis, the researcher would be able to examine whether particular regression slopes in each school vary systematically across schools.[3]

Only in the final step (i.e., the Level 2 HLM model) do we estimate school effects. At this level, the outcome, adjusted for student characteristics included in the Level 1 model, is explored as a function of school characteristics. These characteristics include measures of school structure (e.g., whether the school is private or public, school size, whether the school is located in an inner city, a suburb, or a rural area), school social composition (e.g., average SES, the proportion of minority students, the proportion of immigrant students), school academic organization (e.g., the structure of the curriculum, tracking structure, graduation requirements), and school social organization (e.g., attitudes of teachers toward students, the principal's leadership style, school communal organization). As in Level 1, we may introduce school characteristics in Level 2 both for the purpose of statistical control and as school features that constitute the particular focus of the study. In Level 2, we may explore two types of outcomes as functions of school characteristics: average outcomes (e.g., mean achievement) or social distribution parameters (e.g., the relation between SES and achievement or race differences in each school).[4]

Expanded uses of HLM for contextual effects. HLM may be used for exploration of contextual effects other than in schools. The models described previously, and the examples that follow, are somewhat simplified to give readers a general notion of what school-effects analyses using HLM look like. Depending on the structure of the available data and the questions posed, it is possible to explore three-level HLM models in a school-effects analysis. For example, we may have data on students nested in classrooms, which are also nested in schools. Level 1 would be students, Level 2 would be classrooms, and Level 3 would be schools. We may also be interested in how school characteristics influence achievement growth (i.e., change over time in achievement). In that instance, Level 1 involves a within-student model in which change over time on the same outcome is investigated (e.g., achievement growth in mathematics over the elementary school years). At Level 2, in this context, we would investigate how achievement growth varies as a function of characteristics of students. Level 3 would explore how achievement growth varies as a function of schools, taking characteristics of students into account.

I mention these possibilities so that readers may recognize that with HLM, the exploration of school contexts within a school-effects framework may be more complex than what I have described. However, in the remainder of this article, I restrict the examples to two-level HLM school-effects studies.

EXPLORING SCHOOL EFFECTS: STUDY 1. SCHOOL STRUCTURE

A Focus on School Size

Two strands of research. A defining characteristic of any school is the number of students for which it holds responsibility for educating. Particularly in high schools, enrollment size has important implications for how the school is organized, the curriculum that is offered, and how school members interact. In the study summarized here, Julia Smith and I (Lee & Smith, 1997) considered how this important feature of a high school's structure—the number of students who attend—influences learning in two core subjects: mathematics and reading comprehension.

Much research on this subject couches the size question as, "Is bigger better than smaller?" However, Smith and I (Lee & Smith, 1997) posed more practical questions with more direct policy implications. What is the ideal size of a high school? Does an ideal school size, defined in terms of maximal learning, also support an equitable social distribution of learning? Do size effects vary by the types of students enrolled in the school? Research on school size has followed two separate strands. In one strand, size is viewed in terms of economic efficiency and curriculum specialization. In the second strand, size is seen in social terms, typically defined with a focus on school communal organization. Research in the first strand typically favors larger high schools, whereas conclusions from research framed in the second strand favors small schools.

The second strand of research concludes that students benefit in several ways from attending smaller schools. A more constrained curriculum in many small high schools is typi-

[3]Decisions about how to treat Level 1 variables depend on their use in the HLM model. Statistical controls are typically fixed and centered on the mean for the entire sample. That is, they are standardized around the entire sample mean. Social distribution parameters are left free, and they are centered on their respective school means. That is, such variables are standardized separately for each school. Bryk and Raudenbush (1992) provided more detail about centering.

[4]It may seem that these models ignore an obvious nesting level in educational research—the classroom. After all, students are nested in classrooms, which are then nested in schools. There are two reasons why, in these analyses, the classroom as an explicit level is not taken into account. One is the structure of high schools. We are measuring learning in reading and mathematics over a series of years. However, during these years, students' experiences in English and mathematics classrooms are quite varied (e.g., sometimes they have two different classes per year). Thus, the classroom nesting is quite difficult to conceptualize. The second reason for ignoring the classroom is the nature of the National Education Longitudinal Study of 1988 (NELS:88) data structure. The within-school samples for both studies consist of fewer than 15 students per school. It is unlikely that many of these students experienced their English and mathematics instruction in the same classes.

cally composed of academic courses, so that almost all students follow the same course of study regardless of their interests, abilities, or social backgrounds. This constraint results in both higher average achievement and achievement that is more equitably distributed (Lee & Bryk, 1988, 1989). Social relationships are also more positive in smaller schools. The preponderance of sociological evidence about high schools suggests that smaller is better (Lee et al., 1993).

Refining the question. Although one line of research recommends larger schools as a move toward economic efficiency, these studies focused mostly on school consolidation in rural areas. Most recent research, and much activity in the contemporary policy area aimed at high school reform, recommends that most secondary schools should be smaller than they are. However, there is some logic to support a concern that high schools could become too small to serve their students well. Unless the school population is very homogeneous in terms of ability and aspirations, it is difficult to imagine that a very small high school could offer a curriculum sufficient to meet the needs of an academically diverse student body. We tried to identify the optimal size for a high school. As we defined it, optimal was indicated by how much students learn over their 4 years of high school in two subjects: mathematics and reading comprehension.

Results of Study 1: School Size and Learning

Sample and measures. The samples for these analyses included 9,812 Grade 12 students in 1992 who attended 789 public, Catholic, and elite private high schools, drawn from three waves of data in a nationally representative sample from the National Education Longitudinal Study (NELS:88; Ingels et al., 1994). All students remained in the same high schools between 10th and 12th grade.

We evaluated the effects of school size on learning in mathematics and reading. Our indicator of learning was the change in achievement or gain in those subjects (on equated tests) from the 8th grade to the end of high school (i.e., 12th grade). We chose these subjects because skill in these areas is generally agreed to be important for students' future success. As the content in these subjects is distinct, school size could influence learning in the two quite differently.

The major independent variable was school enrollment size. Because the continuous measure of school size was not normally distributed but rather was strongly skewed in a negative direction, we chose to use categories representing actual enrollments. To decide on these categories, we conducted sensitivity analyses for both learning and its social distribution, by SES and minority status, in each subject. From these analyses, we learned that the effects of size on learning are not linear. To accommodate these nonlinear effects, we converted the continuous measure of school size into a set of eight categories in groups of 300 students: 300 or less; 301 to 600; 601 to 900; 901 to 1,200; 1,201 to 1,500; 1,501 to 1,800; 1,801 to 2,100; and more than 2,100.

We know from other research that many additional characteristics of students and schools influence learning. To eliminate the possibility of alternative explanations for our findings about school size, we took into account some of these factors. Students' gender, SES, minority status, and ability as they entered high school are the characteristics we included at Level 1 of our HLM analyses. We also controlled for several school characteristics at Level 2: school average SES, school minority concentration, and school sector (i.e., Catholic schools and elite private schools were each contrasted with public schools). Our dependent variables—achievement gain in mathematics and reading over the 4 years of high school—were standardized ($M = 0$, $SD = 1$). Thus, our results are presented in effect size (standard deviation) units and may be compared across the two subjects.

Analysis strategy and presentation. We used HLM to evaluate the effect of school size on student learning. Not only did we aim to identify the ideal size of a high school for learning, but we also wanted to identify the particular school size category that was associated with more social equity in learning. Therefore, we also investigated size effects on the SES–learning slope estimated in each school. A third question motivating our analyses was whether school size effects varied by the demographic characteristics of students attending the schools.

The hierarchical models for this study are quite complex. To simplify the presentation of multivariate and multilevel results, we display them here in graphic form. That means that the effects of the student- and school-level covariates are not shown, although we took them into account. An exception is that we use a numerical format to indicate descriptive information about students and schools by size categories. Readers interested in more details are referred to Lee and Smith (1997).

Descriptions of students and schools by enrollment size. What kinds of students attend schools in the eight enrollment size groupings into which we have divided our almost 9,000 students and 800 schools? What school characteristics are associated with size? Table 1 presents descriptive information about students and schools separately by size category. Mean differences between groups are not tested for statistical significance; the information in Table 1 is meant to be descriptive only. The median-size school enrolls about 1,200 students. Although there were quite a few small schools in this sample, more students attended large schools. Because the original NELS sampling strategy called for drawing a close-to-fixed number of students in each NELS school, both student and school sample sizes were reasonably well distrib-

TABLE 1
Study 1: Characteristics of Students and Schools for Several Categories of High School Size

School Size	*Below 300*	*301–600*	*601–900*	*901–1,200*	*1,201–1,500*	*1,501–1,800*	*1,801–2,100*	*Over 2,100*
Student sample size	912	830	1,667	1,645	1,319	1,205	1,263	971
Variables describing students								
Outcomes								
Math gain	8.91	12.13	15.69	13.44	12.20	11.61	10.18	7.84
Reading gain	4.54	6.28	7.61	6.46	5.05	4.60	4.34	3.45
Control variables								
Ability, math[a]	0.03	0.17	0.17	0.18	0.12	0.18	0.05	0.11
Ability, reading[a]	0.05	0.21	0.14	0.19	0.13	0.21	0.07	0.15
Percentage of girls	52.80	51.50	47.90	49.90	52.70	52.40	52.90	50.40
Percentage of minorities	14.50	24.30	14.30	18.00	16.60	15.60	23.50	21.50
Social class[b]	–0.12	0.07	0.11	0.05	0.03	0.08	–0.04	–0.06
Variables describing schools								
School sample size	75	67	148	139	83	70	101	106
Average SES[b]	–0.21	0.09	0.18	0.08	0.09	0.18	–0.15	–0.32
Percentage of high-minority[c] schools	20.30	26.90	16.30	21.20	15.80	14.50	26.10	33.30
Percentage of public schools	95.00	92.50	75.50	81.20	90.80	89.40	92.80	95.90
Percentage of Catholic schools	2.50	4.50	10.90	12.20	6.60	6.60	0.90	3.10
Percentage of independent schools	2.50	3.00	13.60	6.60	2.60	4.00	6.30	1.00

Note. SES = socioeconomic status.
[a]Students' average achievement at eighth grade in the three other subjects was used as a proxy measure of ability, $M = 0$, $SD = 1$. [b]Variables are z scored at $M = 0$, $SD = 1$ on this sample. [c]Schools with more than 40% minority students (Black or Hispanic) are coded 1, others are coded 0, due to a nonnormal distribution.

uted across the groups. In general, learning gains are largest in moderate-size to smaller high schools, although not in the smallest ones. However, schools of these sizes enroll somewhat more able and higher SES students. Other school factors vary considerably by size. For example, average SES is higher in the moderate-size schools and lowest in the largest schools. Minority concentrations are also higher in the largest schools. Many private schools are moderately small.

Size effects on learning: Effectiveness. The differences in Table 1 indicate the importance of taking other characteristics of students and schools into account in analyses that aim to estimate the unique effect of school size on student learning. Figure 1 displays the effects of school size on learning in mathematics and reading. Effects are compared to the size category that contains the most students: 1,201 to 1,500 students. Thus, by definition, the effects in this category are zero. As mentioned, the findings are reported in effect size units.[5] Because our choice of contrasts here is rather arbitrary, the discussion concentrates on the relative magnitudes of school size effects rather than the statistical significance of group differences.

School size shows a similar effect pattern on learning in both subjects, although effects are larger on learning in mathematics (i.e., the dark bars) than reading (i.e., the light bars). Students in moderately sized schools, particularly in the 600 to 900 range, learn most. Although learning is low in the smallest schools (i.e., under 300 students) for both subjects, gains are lowest in the largest schools (i.e., over 2,100 students). These results have been adjusted for students' social background, abilities as they came into high school, and for school social composition and sector. Thus, students learn most in neither the smallest nor the largest high schools. Rather, learning in math and reading is highest in moderate-size high schools, those enrolling 600 to 900 students. Smith and I (Lee & Smith, 1997) concluded that a high school of this size is most effective; effectiveness is measured in terms of average learning in these two core academic subjects.

Size effects on the social distribution of learning: Equity. We (Lee & Smith, 1997) also considered a regression slope as a measure of social equity. More specifically, we explored the relation between students' SES and achievement gains within each school as an indicator of the social distribution of learning. The same HLM analytic models that measure size effects on learning also estimate size effects on the relation between SES and learning in these two subjects. A more equitable social distribution of learning is represented by a less extreme relation of SES to achievement gain in these subjects. Therefore, in the schools in which size categories show negative relations with these outcomes (i.e., where the bars go down instead of up), learning is distributed more equitably. In those results, shown in Figure 2, comparisons are again to schools enrolling 1,200 to 1,500 students.

[5]We follow the lead of Rosenthal and Rosnow (1984) for substantive interpretation of effect sizes. That is, we call effects large if they have a standard deviation of 0.5 or more, moderate if 0.3 through 0.5 , and small if they fall in the 0.1 through 0.3 range. We consider effects smaller than 0.1 *SD* trivial.

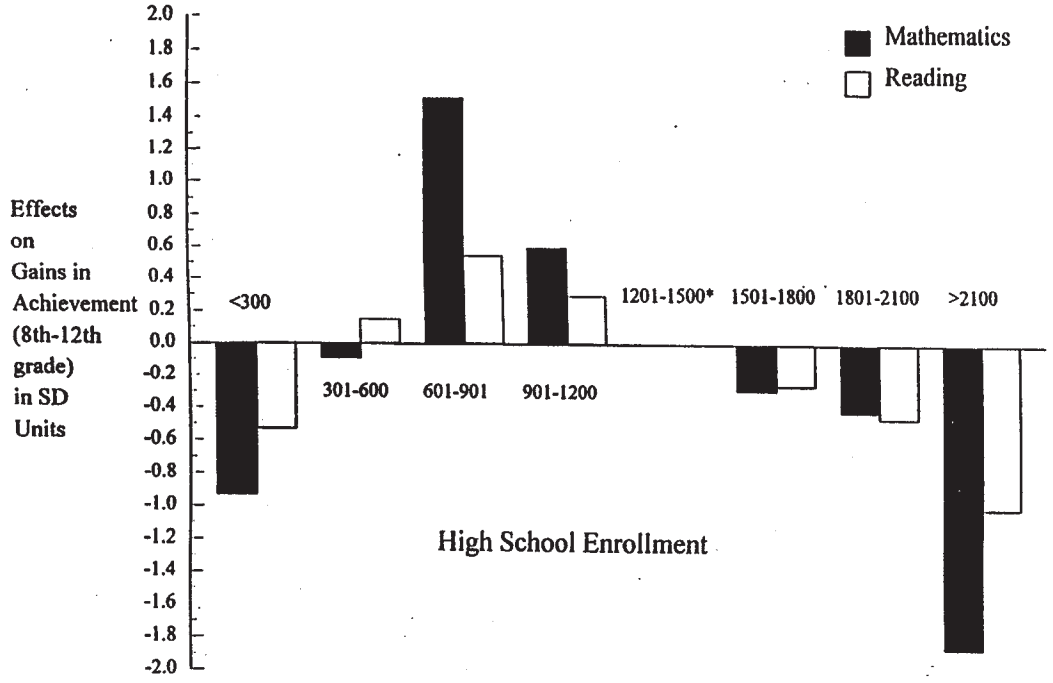

FIGURE 1 Effects of high school size on achievement gains in mathematics and reading. *Note*. 1,201 to 1,500 students were used as the comparison group. By definition, effect sizes were zero in that category.

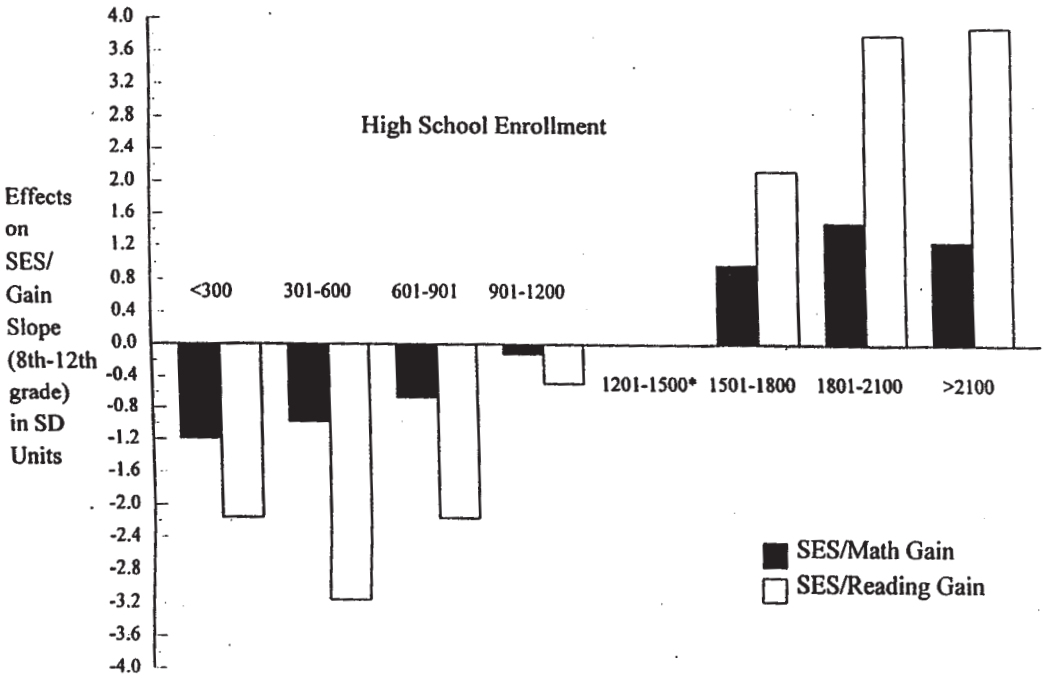

FIGURE 2 Effects of high school size on the relation between SES and achievement gains in mathematics and reading. *Note*. 1,201 to 1,500 students were used as the comparison group. By definition, effect sizes were zero in that category.

Although size effects on average learning (i.e., effectiveness) are larger for mathematics, size effects on the equitable distribution of learning by student SES are larger in reading. Learning is most equitably distributed in smaller schools and least equitably distributed in larger schools. Students' learning of reading is most equitably distributed by SES in schools enrolling 300 to 600 students. Learning in this subject is most inequitably distributed in schools enrolling over 1,800 students. However, learning in mathematics is most equitably distributed in the smallest schools, enrolling 300 students or less. As with reading, larger schools are more inequitable.

To summarize, schools enrolling 600 to 900 students are most effective in terms of average learning. However, even smaller schools with 600 students or less are more equitable. Though size effects on effectiveness are larger in math, size influences equity more strongly in reading. In all cases, our results favor schools that are smaller than most U.S. high schools. The results indicate that schools enrolling fewer than 1,000 students represent contexts that were especially effective and equitable.

Size effects as a function of school composition. We also investigated whether the effects of school size were consistent among schools enrolling different types of students. We identified schools by their social compositions. In one analysis we focused on schools enrolling students of various social class levels; in another, we identified schools enrolling differing proportions of minority students. Results are presented as point-score gains in achievement between 9th and 12th grade rather than as effect sizes. Figure 3 displays the results of our analyses that compare gains in mathematics achievement for students in high-SES and low-SES schools.[6]

Figure 3 suggests three conclusions. The first important, but not surprising, conclusion is that students learn considerably more mathematics over the course of high school in high-SES (i.e., light bars) than in low-SES schools (i.e., dark bars). The second conclusion is that the optimal school size is identical in low- and high-SES schools: 600 to 900 students. The third conclusion—the most striking and most notable—is that school size makes more difference in schools enrolling low-SES students than in those enrolling high-SES students. Most striking are learning differences in the largest high- and low-SES schools. The results shown in Table 1 indicate that school average SES in very large and very small schools is considerably lower than for moderate-size schools. Thus, large numbers of socially disadvantaged students attend schools of a size where learning is lowest. We also investigated whether size is more salient in schools enrolling different proportions of minority students. Size-by-minority concentration interaction terms are statistically significant in both subjects. Similar to findings shown in Figure 2, the optimal school size in both subjects for schools of different social composition is in the middle.

Because of the multilevel nature of this contextual research question—the influence of school size on student learning—this type of analysis could be accomplished only with a multilevel methodology such as HLM. These HLM analyses are quite complex, but the results are quite straightforward. They show that very large high schools are problematic contexts for learning, especially those that enroll large proportions of disadvantaged students. In general, the patterns are quite similar for learning in both subjects. All students, but particularly those who attend high-minority or low-SES schools enrolling over 1,800 students, learn very little. These findings suggest some conclusions about educational policy.

Policy Implications for Study 1

Several conclusions. Our first conclusion is that high schools should be smaller than they are. For several decades, large comprehensive high schools have been the norm in the United States. Economic arguments favored school consolidation, stressing the need for substantial numbers of students to offer a broad array of courses. We contend that these arguments have led the country in the wrong direction in designing its secondary schools. Our results indicate an ideal size for high schools: 600 to 900 students. Although size effects differ somewhat for effectiveness and equity, we conclude that students learn more in smaller high schools.

There seems to be a balance point, however. Schools must be large enough to offer a credible curriculum but small enough that school members can know one another well. Smaller schools function as communities (Barker & Gump, 1964; Bryk et al., 1993; Lee et al., 1993). It seems logical that a school could function well with small size only if the clientele were socially and academically homogeneous, students' aspirations similar, and the resource base high (e.g., elite private schools). However, our results from a large and nationally representative sample of high schools suggest that learning is limited in very small as well as in very large schools. Although the dual aims of effectiveness and equity do not converge on exactly the same school size, we conclude that the ideal size high school should enroll at least 600 students.

The consistency of results across schools defined by differing social compositions is important. Within the full range of high schools in the United States, for schools that enroll mostly high- or mostly low-SES students or schools that enroll high or low proportions of minority students, the ideal size was consistent: 600 to 900 students. The findings suggest

[6] Low-SES schools are those whose average school SES is 1 *SD* below the sample average for school SES, and high-SES schools are 1 *SD* above the sample mean. We created a set of seven size-by-average SES interaction terms and included them in Level 2 in the same HLM model as we used for the analyses in Figures 1 and 2. We also investigated a similar model for achievement gains in reading comprehension. As the interaction terms were not statistically significant as a set, we do not include the results here.

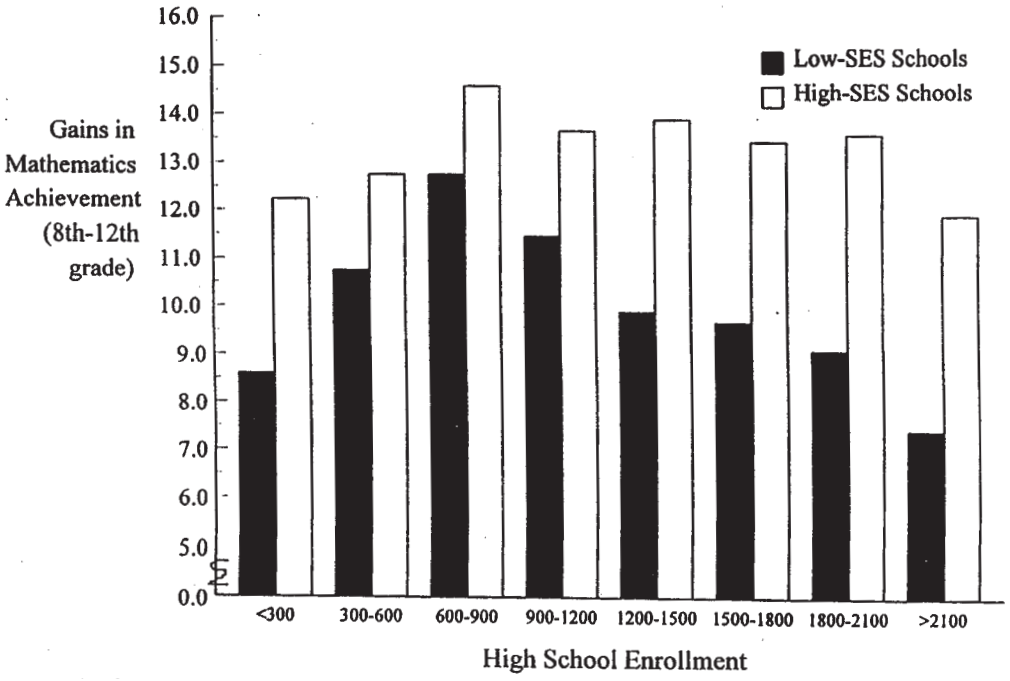

FIGURE 3 Average gains in mathematics achievement by high school size in low-SES and high-SES high schools.

a definitive answer to the question of which size high school works best, and for whom.

Although school size is a significant factor in determining learning for all students, this structural feature has particular salience in school contexts in which disadvantaged students are concentrated. Learning is particularly low in very large high schools, and our results show that in these contexts large numbers of economically and racially disadvantaged students are being educated. Thus, under an aim of inducing social equity in learning for students in secondary schools, it is especially important to consider reducing the size of the largest high schools.

Reconsidering school size. I recognize the difficulty of making U.S. high schools smaller, despite the strength and consistency of the findings of this study. The nation is hardly prepared to invest in building new, smaller schools on a wide scale. Elsewhere, my colleagues and I recommended that policymakers consider the option of dividing large high schools into smaller units (i.e., creating schools within schools; Lee et al., 1993). I am, in fact, engaged in a study of this option at the moment. Preliminary findings suggest that although the idea is logical and school size is currently a hot policy issue, the prevalence of high schools with this design across the country is low (Lee, Ready, & Johnson, 1999).

Although school size is a frequently discussed contextual issue among policymakers and the results shown here are clear and definitive, I suggest that changing school size alone is unlikely to influence student learning directly. Rather, school size probably influences learning indirectly through its association with the academic and social organization of secondary schools (e.g., course offerings, tracking policy, teachers' attitudes, team teaching, collaboration, social interactions among school members). Were we to include good measures of many such features of school organization, it is likely that they would explain away the direct effects of school size on student learning shown in this study. However, it is much easier to change many of these organizational features by reducing the size of a high school. Smaller school size is, then, a necessary but not sufficient policy lever for increasing student learning.

EXPLORING SCHOOL EFFECTS: STUDY 2. SCHOOL SOCIAL ORGANIZATION

A Focus on Teachers' Work

Teachers' attitudes. The main work of schools, teaching and learning, revolves around teachers. Many contemporary school reforms focus on changing procedures, principles, rules, roles, and relationships within schools. Suggestions of these last two types often target teachers and teaching. This study focuses on the professional lives of teachers. Whether the locale is inside or outside the classroom, whether the focus is on the academic or social dimensions of schooling, the work done by teachers is pivotal in the education of students. Moreover, teachers transmit

much more than factual information to their students. The work teachers do (instruction) is influenced by the conditions in which they perform their tasks. A common assumption among reforms that address the work of teachers is that, either implicitly or explicitly, such reforms aim to improve learning for all students.

Study 2 tests this assumption by exploring how characteristics of teachers' work lives influence how much their students learn (Lee & Smith, 1996). Smith and I captured the notion of teachers' professional lives in three ways: how teachers define their work; how they interact with their students, colleagues, and superiors (collaboration); and the degree to which teachers feel control over their work. Because our main focus here is on teachers' attitudes about students, we do not discuss teacher collaboration or control. Readers interested in a full treatment of these subjects, as well as the literature that guided our work, should read the full study. As before, our research question—How do teachers' attitudes about their students influence student learning?—is multilevel. Although the particular context of Study 2 focuses on school social organization rather than school structure as in Study 1, both contexts describe schools. Thus, both require the use of multilevel methods.

Effects of teachers' attitudes on students. Rosenthal and Jacobson's (1968) seminal study of the self-fulfilling prophecy of expectations showed that experimental manipulation of teachers' beliefs about the abilities of students influences learning. Despite the controversy and research that study sparked among psychologists about the effects of teacher expectations on student outcomes, its conclusions have been confirmed (e.g., Brophy, 1983; Cooper, Findley, & Good, 1982; Cooper & Tom, 1984; Raudenbush, 1984; Weinstein, Madison, & Kuklinski, 1995). Teachers' beliefs about students' ability to learn influence achievement. The result suggests that teachers' expectations about their students, as well as their willingness to assume responsibility for the results of their teaching, have important consequences for learners. Students not only become more engaged in school under those circumstances (Firestone & Rosenblum, 1988), but they also learn more (Cooper & Tom, 1984).

Studies of teaching in schools in which tracking and ability grouping are common have suggested that teachers' negative attitudes and expectations about their students' ability to learn potentially distance all but the most able students from high-level learning (Ennis, 1994; Oakes, 1985). Teachers' negative projections can also alienate students, negatively influencing their learning (Firestone & Rosenblum, 1988; Newmann, 1981). In short, evidence supporting the impact of teacher expectations on student learning is compelling. Students fulfill their teachers' prophecies, performing up to, or down to, the projections and standards held for them.

An organizational property of schools. Although much research on this topic has been conducted by psychologists, we locate our research within a sociological framework. The notion of teachers' expectations for students is expanded to a larger context: a school culture that may or may not be centered on expectations. We expand teachers' expectations for learning to their responsibility for learning. The top of this spectrum includes teachers who take personal responsibility for the success or failure of their own instruction. Such teachers view teaching and learning as an interactive process with students cast as active participants, rather than as a one-way flow of information (Brattesani, Weinstein, & Marshall, 1984). Assuming more responsibility for learning also implies that such teachers have greater confidence in their own abilities to affect outcomes of the instructional process (i.e., learning) for all students and not just the most able or motivated. When teachers are responsibile for learning, their expectations about their students' abilities to learn may be linked with their own sense of efficacy in teaching: a personal attitude that teaching is worth the effort. The message of expectancy may be embedded in a teacher's self-efficacy to generate an attitude of organizational commitment or responsibility (Maehr & Braskamp, 1986).

Within a sociological framework, we considered teachers' attitudes in the aggregate, as a feature of the social organization of the school. Students and teachers alike are influenced in their personal commitment to education by the extent to which the faculty in their schools feel that teaching is worth the effort. This collective attitude is part of a larger community dynamic that invests the technical core of schooling with value and worth (Bryk & Driscoll, 1988). Considered in this way, teachers' willingness to take responsibility for their students' learning indicates a collective commitment to caring about students.

Research focus. We investigated the link between high school students' learning and the social organization of their schools, defined in terms of teachers' work. Although our study investigated the collaborative nature of staff relationships and the degree to which teachers see themselves as in control or empowered in their classrooms and in their schools to define these elements of school social organization, the major focus and strongest findings are on school norms reflecting teachers' attitudes about students. This dimension of teachers' work lives, which we call *collective responsibility,* may influence students in two ways. Most obvious, when the teachers of a school as a group believe that their own efforts are crucial in the learning process, student learning could increase. Moreover, when responsibility is assumed for all students, regardless of academic qualifications or social characteristics, learning may be more equitably distributed among the student body. We explore both outcomes.

Results of Study 2: Collective Responsibility and Learning

Samples and measures. The school and student samples for this study are similar to those in Study 1. However, in Study 2 we focus on students' growth in achievement over the first 2 years of high school. Because this sample (also drawn from the first two waves of NELS) contains some students who subsequently dropped out, this sample is somewhat larger (11,692 10th graders in 820 schools). Most schools are public (650), 68 are Catholic, and 47 are elite private schools. Data on teachers' work lives are drawn from teachers' reports in which two teachers are linked with each NELS student (9,904 teachers of English, math, social studies, and science).

The measures of teachers' professional community are factor-score composites constructed from individual teacher reports that were subsequently aggregated to the school level. The measure of collective responsibility for learning includes several related ideas: teachers' internalizing responsibility for the learning of their students, rather than attributing learning difficulties to weak students or deficient home conditions; a belief that teachers can succeed with all students; willingness to alter teaching methods in response to students' difficulties and successes; and self-efficacy in teaching.[7] Other work life measures tap cooperation among teachers and the degree of control teachers feel they have over teaching and school policy. We investigate the effects of school social organization on learning in four subjects: math, reading, history, and science. Similar to Study 1, learning is measured as a simple gain between achievement test scores in each subject from the end of the 8th grade to the 10th grade.

To eliminate the possibility of alternative explanations in all our multivariate and multilevel analyses, we took into account other characteristics of students and schools. Student characteristics included gender, SES, minority status, ability, engagement in school in the eighth grade, and high school track placement. We also included controls for the characteristics of schools: school composition (i.e., average school SES, minority concentration), school ability level (i.e., average eighth grade achievement), academic emphasis (i.e., proportion in the academic track), sector (i.e., Catholic and elite private schools compared to public schools), and school size.

Analysis strategy and presentation. We separated the sample schools into three groups: those with high levels of collective responsibility for learning (1 *SD* or more above the mean on that variable), schools with average levels of collective responsibility (within 1 *SD* above or below the mean), and schools with low collective responsibility (1 *SD* or more below the mean). Descriptive information about students and schools is shown separately for these school groupings. Again, the multilevel research question dictated that HLM was the appropriate analysis technique. HLM allows us to investigate not only whether collective responsibility is associated with average student learning in the four subjects but also whether this measure of school social organization is associated with a socially equitable distribution of learning in these subjects. Consistent with the aim of keeping numerical information from these complex analyses to a minimum, we present multivariate results as graphic interpretations from the full-model HLM analyses. As with Study 1, readers interested in the complete numerical results from the several HLM analyses are referred to Lee and Smith (1996). However, the effects shown in the graphs are exact numerical representations of the HLM results. In this case, the multilevel multivariate results are shown as point-score gains in achievement.

Descriptions of students, schools, and collective responsibility. What kind of students attend schools where collective responsibility for learning is quite common, compared to schools where these attitudes are average or unusual? What other school characteristics are associated with schools differentiated by levels of collective responsibility? Such information is basically descriptive. These descriptive questions are addressed by dividing the sample schools into three levels (low, medium, and high) of collective responsibility and examining group means on several variables describing how students and schools differ. This information is presented in Table 2. The criteria used to create these categories led to about two thirds of the schools and three fourths of the students in the average category.

Students attending schools with high levels of collective responsibility are advantaged both socially and academically compared to students in schools where collective responsibility is uncommon (see "Variables describing students" in Table 2). This pattern is reflected in achievement gains in all subjects. Schools with high levels of collective responsibility enroll students from families of higher SES. Students attending schools typified by high levels of collective responsibility

[7] There were 12 NELS items on the teacher survey we used to construct our composite with principle components factor analysis. Items are listed in order of their factor weights. Each had a 5-level response set ranging from 1 (*strongly disagree*) to 5 (*strongly agree*):

- Little I can do to insure high achievement (reversed).
- I can get through to the most difficult student.
- Different methods can affect a student's achievement.
- Teachers make a difference in students' lives.
- It is a waste of time to do my best in teaching (reversed).
- Teachers are responsible for keeping students from dropping out.
- Students' attitudes reduce academic success (reversed).
- I work to create lessons that students will enjoy learning.
- Students' success or failure is due to factors beyond me (reversed).
- Students are incapable of learning the material (reversed).
- I change my approach if students aren't doing well.
- Student misbehavior interferes with my teaching (reversed).

TABLE 2
Study 2: Characteristics of Students and Schools for Three Levels of Collective Responsibility for Learning

	Level of Collective Responsibility for Learning		
	High	*Medium*	*Low*[a]
Student sample size	1,226	8,801	1,665
Variables describing students			
Outcomes			
Math gain	6.57	5.39	4.95
Reading gain	3.70	2.51	1.61
History gain	2.95	1.51	1.26
Science gain	3.43	1.54	1.33
Student controls			
Engagement (8th)[b]	0.24	–0.01	–0.12
Social class[b]	0.46	–0.03	–0.21
Minority status	0.16	0.20	0.18
Female status	0.50	0.51	0.51
Ability controls (8th grade):			
Math control[b]	0.45	0.05	–0.10
Reading control[b]	0.50	0.06	–0.11
History control[b]	0.50	0.07	–0.11
Science control[b]	0.51	0.06	–0.11
Curriculum track (10th grade):			
Academic track	0.42	0.38	0.34
General track	0.49	0.50	0.51
Vocational track	0.09	0.12	0.15
Variables describing schools			
School sample size	134	548	138
Characteristics of professional community			
Average collective responsibility[b]	0.87	0.02	–0.49
Variability in collective responsibility	0.47	0.48	0.61
Cooperation[b]	0.66	–0.05	–0.66
Control[b]	0.80	–0.04	–0.48
School controls			
Average SES[b]	0.77	–.083	–0.37
Percentage of minority students	21.54	24.280	23.98
Percentage in academic track	70.46	48.260	33.15
Average achievement in Grade 8	55.22	50.980	49.83
School size	1,077	1,280	844
School sector			
Percentage of public schools	44.05	91.90	98.86
Percentage of Catholic schools	21.43	4.54	0.00
Percentage of independent schools	34.52	3.57	1.14

Note. SES = socioeconomic status.
[a]School categories created as follows: Low = < 1 *SD* below sample mean on collective responsibility; Medium = between –1 *SD* and +1 *SD* from the mean; High = > 1 SD above the sample mean. [b]Variables are *z* scored, $M = 0$, $SD = 1$.

are also more able in all subjects, more engaged in school, and more likely to be in the academic, and least likely to be the vocational track. Schools of average collective responsibility enroll slightly more minority students than the schools at either extreme. Most group mean differences are statistically significant.

The pattern of differences among students is generally reflected in schools, except that group differences were typically even larger (see "Variables describing schools" in Table 2). Schools with high collective responsibility were also characterized by more cooperation and more teacher control, suggesting that the measures of teacher professional community are related to each other. More variability among teachers' responsibility for learning in schools with lower collective responsibility means that teachers share these attitudes less in low-responsibility schools. Group differences in average SES were large. Average SES of high-responsibility schools differed from their low-responsibility counterparts by 1.4 *SD*. Such schools also enrolled students with higher average ability. Compared to the middle- and low-responsibility schools, the high proportion of private schools (both Catholic and independent) in the high-responsibility school group is noteworthy. Patterns for school size show that the largest schools are in the middle category; schools with low collective responsibility are the smallest. Most group differences among school characteristics were statistically significant.

The results shown in Table 2 suggest that schools with high collective responsibility enroll more socially and academically advantaged students. Collective responsibility is also associated with school sector, with private schools considerably higher on this measure. It is clear that these social and academic characteristics of students and compositional and structural characteristics of schools should be taken into account in analyses aiming to isolate the effect of teacher professional community on learning.

The Effects of Professional Community on Student Learning

Between-school proportions of variance in the outcomes. As described previously, school-effects HLM analyses are typically performed in three steps. In the first step, variance in the outcome measures is partitioned into its within- and between-school components. Gains in achievement in mathematics, reading, history, and science in the first 2 years of high school represent our outcomes. The proportion of the total variance in these measures that lay between schools ranged from 15% to 20%, lower for reading and higher for science. Our search for effects on student learning of teacher professional community involves only the variance between schools.

Taking students' demographic and social characteristics into account. The second step in these HLM analyses estimated within-school HLM models in which we introduced the controls for student social and academic background described previously. Virtually all measures of students' demographic and academic background were related strongly to achievement gains in the four subjects. Academic (compared to general) track placement was strongly and positively associated with achievement gain, whereas vocational track placement was negatively related to achievement gains.

We gave special focus in our within-school HLM models to the relation between SES and achievement gains. In these four core subjects, SES was significantly and positively related to gains, an indication that learning was inequitably distributed. Within U.S. secondary schools, students from families of higher social class learned more math, reading, history, and science than their less advantaged classmates. We investigated the relation between SES and achievement gains as outcomes. Thus, our between-school HLM models investigated eight outcome measures: school average achievement gains in math, reading, history, and science and the relation between SES and achievement gain in each subject. Each outcome was adjusted for the students' demographic (i.e., gender, minority status) and academic status (i.e., ability, engagement, and track placement) within each school.

Multilevel analyses that take school composition and structure into account. The full HLM models investigated the effects of school professional community on average gains in achievement and also on the relation between SES and achievement gain. Similar to Study 1, these two types of outcomes measured effectiveness and equity. Each HLM analysis, for achievement in the four subjects, included all controls for students' academic and social background, school composition, and school structure described previously. They also included all measures of teacher professional community.

Among the four measures of school professional community, only one—responsibility for learning—was consistently, positively, and significantly related to achievement gain. Moreover, variability within schools on this measure was negatively associated with achievement gains in math and reading. This finding suggests that when willingness to take responsibility for students' learning is not a common attitude among the faculty, students learn less. Taken together, these results indicate that students learned more in schools where teachers consistently reported themselves willing to assume responsibility for their students' learning. The other measures of teacher professional community, however, had few effects. As we did in Study 1, we also explored social inequality as the relation between SES and learning in each school. We found that responsibility for learning was consistently, significantly, and negatively related to the social distribution of learning in all subjects.

In sum, we found that in schools in which collective responsibility was high, students learned more; learning was less related to their social background. Effects of the other measures of school professional community (i.e., collaboration, control, and the distribution of collective responsibility) on learning and its equitable distribution were either inconsistent, small, or statistically insignificant. Rather than concluding that these features of teacher professional community are unimportant, I believe that these findings resulted from the intercorrelation among these measures. Descriptive results from Table 2 support this suggestion.

Revisiting collective responsibility for learning. These relations may be easier to understand if they are displayed graphically. Figure 4 displays achievement gains in

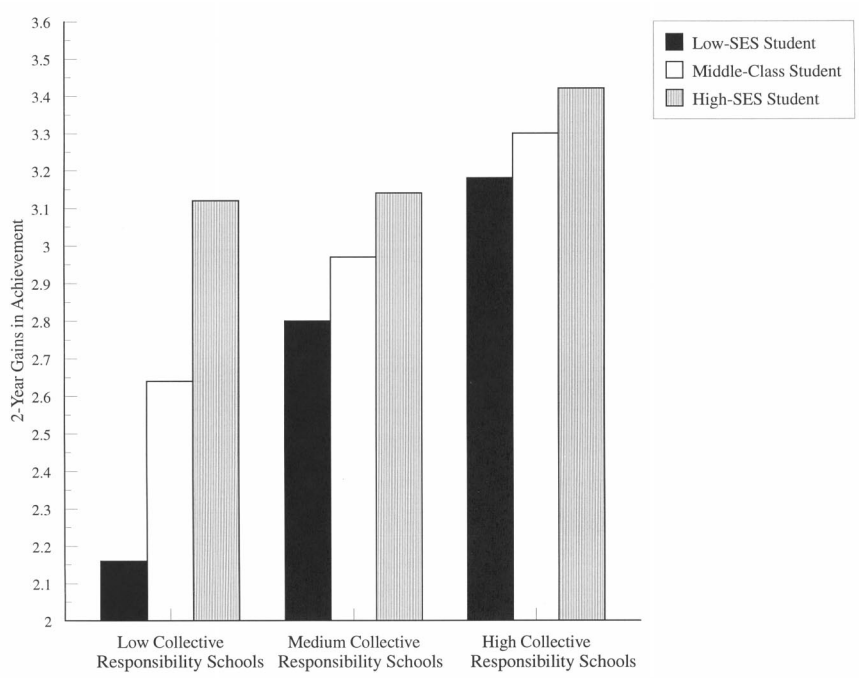

FIGURE 4 Achievement gains in mathematics for students of varying levels of SES in schools with different levels of collective responsibility.

mathematics for nine groups: students from low-SES, middle-class, and high-SES families who attend schools whose levels of collective responsibility are low, medium, and high. Low-SES students are 1 SD below the mean on that variable (i.e., the dark bars); middle-class students are those whose SES value is the mean (i.e., the white bars); high-SES students are those whose SES is 1 SD above the mean (i.e., the gray bars). Achievement gains for these three groups of students are related to the schools they attend, with the schools differentiated by average collective responsibility. These levels are constructed in the same way as the three groups described in Table 2. The heights of the bars in this graph provide a visual representation of the numerical results in the HLM analyses. The results and the method we used to compute these graphs may be found in Lee and Smith (1996).

Figure 4 displays achievement gains for students in mathematics from 8th to 10th grade. Three trends are evident: (a) Students with higher SES learn more in math than their lower SES peers; (b) students in schools in which collective responsibility is high learn more than their counterparts in schools with low collective responsibility; and (c) there are larger differences in achievement gains among low-, middle-, and high-SES students in schools with low collective responsibility than in schools where the level is high (i.e., SES has less of an influence on how much students learn in mathematics in schools where teachers take more responsibility for their students' learning). High-SES students learn more in all types of schools, but students' math learning, especially students from low-SES families, is quite strongly affected by the type of schools they attend. When low-SES students attend schools in which most teachers take responsibility for student learning, they learn quite a lot in the early years of high school. As the patterns are quite similar for learning in the three other core subjects, the results are not included here.

Policy Implications for Study 2

How do we change teachers' attitudes about their students? In one sense, the strength and consistency of these findings are troubling, in that high schools vary markedly on collective responsibility. In many U.S. high schools, particularly those that enroll many students who are academically and socially disadvantaged, teachers' beliefs about the limitations of their students' ability to learn and of teachers' ability to teach them effectively are quite prevalent. I suggest that high school teachers should assume responsibility for their students' academic progress, along with the students themselves. The fact is that there is much variation in this measure across the high schools in the nation. Moreover, the fact that schools in which teachers are less willing to assume such responsibility to enroll large proportions of disadvantaged students is vexing. Although these results demonstrate how important teachers' attitudes are to students' learning, how to change these attitudes is not obvious.

These findings are consistent with research that has demonstrated that cooperative, collegial, and communal school environments have strong effects on student engagement and teacher commitment (Bryk & Driscoll, 1988; Bryk et al., 1993). Although one would expect that the effects of teachers' attitudes on their students' learning would be linked primarily through the instructional process (which we did not explore), the results suggest that they also influence students through their effects on the social organization of schools.

Professional community is important. Calls for reorganizing schools from the Center for Organization and Restructuring of Schools (CORS) included a focus on the professional life of teachers. Research sponsored by CORS advocates development of a professional teaching community (Louis, Marks, & Kruse, 1996; Newmann & Associates, 1996). Several elements of professional community, as those authors defined it, are captured in Study 2.[8] Louis et al. listed collective focus on student learning and collaboration as two of five critical elements defining a professional community. Among several structural elements necessary for this type of community to develop and grow within schools, they listed teacher empowerment and school autonomy. The findings of Study 2 offer considerable empirical support for changing the work lives of teachers in secondary schools in the direction laid out by CORS. Schools in which teachers interact as a professional community have benefits for their students, both for learning and for social equity.

DISCUSSION: THE SCHOOL CONTEXT

Reviewing the Structure of School Effects Studies

Two examples. The studies described in this article were offered as examples of research that investigates the specific educational context defined by the school. Both studies explored how characteristics of secondary schools influence the students who attend them. The particular characteristics of secondary schools explored here include a feature of school structure, namely the number of students the school enrolls, and a feature of school social organization, namely teachers' collective attitudes about their responsibility for learning. Both studies use large, current, longitudinal, and nationally representative samples of high schools and their students. Although these studies do not employ an experimental design (and it is hard to imagine how students could be randomly assigned to schools of different size or different social organization), a major strength is their longitudinal designs. Learning was defined by the points students gained on tests

[8] I was a member of the CORS research team, which sponsored Study 2.

administered before entering high school and 2 or 4 years later. Both studies focus on academic outcomes, defined in terms of both effectiveness (i.e., average learning in particular subjects) and equity (i.e., the social distribution of learning by student SES).

Implications for policies about the school context. Findings from these studies have implications for social policies about how our nation defines the contexts in which its adolescent citizens are educated. The findings from Study 1 suggest that U.S. students learn more in mathematics and reading in high schools that are typically smaller than schools most students attend, but not too small. The study also found that high school size is more important for socially disadvantaged students than for their more affluent counterparts. Drawing direct policy implications from the findings of Study 2 is more difficult. Findings indicate that teachers' attitudes about their students—in particular, teachers' willingness to take personal responsibility for their students' learning—influence learning in several subjects. Moreover, in schools in which teachers collectively take such responsibility, all students seem to learn. That is, such schools are both more effective and more equitable. Although these findings give us a guide for how teachers' attitudes influence how much students learn, it is less clear how to change teachers' attitudes, at least through direct policy intervention.

How Do We Study Educational Contexts?

Appropriate methods. I had several purposes in describing these two studies. One was to demonstrate that the school context is important. The strength and consistency of their findings should convince readers that how schools are structured and organized can affect the students who attend them. A second purpose was to provide some information about and examples of a methodology that is most appropriate for conducting studies of educational contexts, namely HLM. I encourage researchers who are interested in the study of educational contexts to consider learning about and using HLM in their studies of context. I argue that contextual research questions are, by their very nature, multilevel (i.e., individuals are nested in these contexts). I contend that it is inappropriate to analyze multilevel data and address multilevel research questions without employing multilevel methods.

Appropriate data. The use of HLM to investigate context effects, however, requires access to data with multilevel structures. That is, data are required on several individuals within each context, and there must also be rich descriptions available about the contexts themselves, beyond the aggregated characteristics of individuals. At the very least, contextual researchers in the field of education who are interested in quantitative analysis need to design their data collections appropriately; typically, this requires a substantial number of individual cases per context, measures of individuals' outcomes and perhaps treatments over time, and a substantial number of contexts—often a costly undertaking.

Researchers with these interests may want to consider becoming familiar with data collected by and available from the National Center for Education Statistics (NCES). Through its Longitudinal and Household Studies branch, NCES routinely collects data with multilevel structures. Although the most recent longitudinal studies from NCES—NELS:88 and High School and Beyond—focused on U.S. students' experiences in public and private high schools, the first two waves of a new NCES longitudinal study, the Early Childhood Longitudinal Study (ECLS), will be available in mid 2000. ECLS focuses on young children, beginning in kindergarten. These children will be followed through elementary school, with extensive data available from parents, teachers, and schools. Thus, a third purpose is to encourage researchers to consider the structure of the data they collect or use. Contextual research questions are, by their nature, multilevel. Thus, researchers need to design data collections and employ methods that are most appropriate for studying educational contexts in general and school contexts in particular.

Why Should We Study Educational Contexts?

I hope that readers of the articles in this issue are convinced of the importance of studying educational contexts. Children's learning is srongly influenced by the contexts in which it occurs. Those contexts may be defined by the children's families, the classmates with whom they experience schooling, the peers with whom they choose to interact, and the teachers who instruct them. With this article I have argued that students are also profoundly influenced by the schools they attend.

The arguments made here for the importance of considering the context of learning are not meant to downplay the importance of children's own inputs to learning. Beyond their ability or prior achievement, children's attitudes and behaviors have a strong influence on their learning. Although in this article I did not explore school effects on outcomes that measure such attitudes and behaviors as engagement, commitment, motivation, attendance, working hard, doing homework, or choosing challenging courses, there is much evidence that schools have profound effects on school-related attitudes and behaviors, as well as learning and its social distribution. What-

ever the outcomes explored, the contexts in which children experience their education influence their educational progress.

REFERENCES

Barker, R., & Gump, R. (1964). *Big school, small school: High school size and student behavior.* Stanford, CA: Stanford University Press.

Bidwell, C. (1965). The school as a formal organization. In J. G. March (Ed.), *Handbook of organizations* (pp. 972–1019). Chicago: Rand McNally.

Brattesani, K. A., Weinstein, R. S., & Marshall, H. H. (1984). Student perceptions of differential teacher treatment as moderators of teacher expectation effects. *Journal of Educational Psychology, 76,* 236–247.

Brookover, W. B., Beady, C., Flook, P., Schweitzer, J., & Wisenbaker, J. (1979). *School social systems and student achievement: Schools can make a difference.* New York: Praeger.

Brophy, J. E. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology, 75,* 631–661.

Bryk, A. S., & Driscoll, M. E. (1988). *The school as community: Theoretical foundations, contextual influences, and consequences for students and teachers.* Madison: University of Wisconsin, National Center on Effective Secondary Schools.

Bryk, A. S., Lee, V. E., & Holland, P. B. (1993). *Catholic schools and the common good.* Cambridge, MA: Harvard University Press.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage.

Clark, D. L., Lotto, L. S., & McCarthy, M. M. (1980, March). Factors associated with success in urban elementary schools. *Phi Delta Kappan,* 469–470.

Coleman, J. S., & Hoffer, T. (1987). *Public and private high schools: The impact of communities.* New York: Basic Books.

Coleman, J. S., Hoffer, T., & Kilgore, S. B. (1982). *High school achievement: Public, Catholic and private schools compared.* New York: Basic Books.

Cooper, H. M., Findley, M., & Good, T. (1982). Relations between student achievement and various indexes of teacher expectations. *Journal of Educational Psychology, 74,* 577–579.

Cooper, H. M., & Tom, D. Y. (1984). Teacher expectation research: A review with implications for classroom instruction. *Elementary School Journal, 85,* 77–89.

Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership, 37*(1), 15–24.

Ennis, C. D. (1994). Urban secondary teachers' value orientations: Social goals for teaching. *Teaching and Teacher Education, 10,* 109–120.

Firestone, W. A., & Rosenblum, S. (1988). Building commitment in urban schools. *Educational Evaluation and Policy Analysis, 10,* 285–300.

Greeley, A. M. (1982). *Catholic high schools and minority students.* New Brunswick, NJ: Transaction Books.

Hoffer, T., Greeley, A. M., & Coleman, J. S. (1985). Achievement growth in public and catholic schools. *Sociology of Education, 58,* 74–97.

Ingels, S. J., Dowd, K. L., Baldridge, J. D., Stipe, J. L., Bartot, V. H., & Frankel, M. R. (1994). *National Education Longitudinal Study of 1988: User's manual* (Report No. NCES 94–374). Washington, DC: U.S. Department of Education, Office of Education Research and Improvement.

Lee, V. E., & Bryk, A. S. (1988). Curriculum tracking as mediating the social distribution of high school achievement. *Sociology of Education, 61, 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,* 171–267. Washington, DC: American Educational Research Association.

Lee, V. E., Ready, D. D., & Johnson, D. J. (1999, August). *High schools divided into schools-within-schools: Prevalence and design formats.* Paper presented at the 1999 annual meeting of the American Sociological Association, Chicago.

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,* 103–145.

Lee V. E., & Smith, J. B. (1997). High school size: Which works best and for whom? *Educational Evaluation and Policy Analysis, 19,* 205–227.

Lesko, N. (1988). *Symbolizing society.* Philadelphia: Falmer.

Louis, K. S., Marks, H. M., & Kruse, S. (1996). Teachers' professional community in restructuring schools. *American Educational Research Journal, 33,* 757–798.

Maehr, M. L., & Braskamp, L. A. (1986). *The motivation factor: A theory of personal investment.* Lexington, MA: Heath.

Merton, R. K. (1949). *Social theory and social structure: Toward a codification of theory and research.* Glencoe, IL: Free Press.

Newmann, F. M. (1981). Reducing student alienation in high schools: Implications of theory. *Harvard Educational Review, 51,* 546–564.

Newmann, F. M., & Associates. (1996). *Authentic achievement: Why are some schools more successful than others?* San Francisco: Jossey-Bass.

Oakes, J. (1985). *Keeping track: How schools structure inequality.* New Haven, CT: Yale University Press.

Parsons, T. (1951). *The social system.* Glencoe, IL: Free Press.

Purkey, S. C., & Smith, M. S. (1983). Effective schools: A review. *Elementary School Journal, 83,* 427–454.

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,* 85–97.

Rosenholtz, S. J. (1985). Effective schools: Interpreting the evidence. *American Journal of Education, 93,* 352–388.

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom.* New York: Holt, Rinehart & Winston.

Rosenthal, R., & Rosnow, R. L. (1984). *Essentials of behavioral research: Methods and data analysis.* New York: McGraw-Hill.

Rutter, M., Maughan, B., Mortimore, P., Outson, J., & Smith, A. (1979). *Fifteen-thousand hours: Secondary schools and their effects on children.* Cambridge, MA: Harvard University Press.

Sorokin, P. (1928). *Contemporary sociology theories.* New York: Harper & Brothers.

Waller, W. (1932). *The sociology of teaching.* New York: Russell & Russell.

Weber, M. (1947). *Theory of social and economic organization* (A. M. Henderson & T. Parsons, Trans.). New York: MacMillan. (Original work published 1926)

Weinstein, R. S., Madison, S. M., & Kuklinski, M. R. (1995). Raising expectations in schooling: Obstacles and opportunities for change. *American Educational Research Journal, 32,* 121–159.