Schools and Inequality: A Multilevel Analysis of Coleman's Equality of Educational Opportunity Data

GEOFFREY D. BORMAN
MARITZA DOWLING

University of Wisconsin–Madison

Background/Context: *The Equality of Educational Opportunity study is widely recognized as one of the most important studies on schooling ever performed. The findings from the report have shaped the field of education, national education policies, and wider public and scholarly opinion regarding the contributions of schools and schooling to equality and productivity in the United States. Despite past reanalyses of the data and decades of research on the effects of schools as organizations, the report's fundamental finding—that a student's family background is far more important than school social composition and school resources for understanding student outcomes—still retains much of its currency. Purpose/Objective/Research Question/Focus of Study: Using the original Equality of Educational Opportunity data, this study replicated Coleman's statistical models but also applied a two-level hierarchical linear model (HLM) to measure the effects of school-level social composition, resources, teacher characteristics, and peer characteristics on ninth-grade students' verbal achievement.*

Research Design: *HLM allows researchers to disentangle how schools and students' family backgrounds contribute to learning outcomes. The methodology offers a clearer interpretation of the relative effects of school characteristics, including racial/ethnic composition, and family background, including race/ethnicity and social class, on students' academic outcomes.*

Findings/Results: *Our results suggest that schools do indeed matter, in that when one examines the outcomes across the national sample of schools, fully 40% of the differences in achievement can be found between schools. Even after statistically taking into account students' family background, a large proportion of the variation among true school means is related to differences explained by school characteristics. Within-school inequalities in the*

Teachers College Record Volume 112, Number 5, May 2010, pp. 1201–1246

0161-4681

achievement outcomes for African American and White students and students from families of higher and lower social class are explained in part by teachers' biases favoring middle-class students and by schools' greater reliance on curriculum differentiation through the use of academic and nonacademic tracking.

Conclusions/Recommendations: *Formal decomposition of the variance attributable to individual background and the social composition of the schools suggests that going to a high-poverty school or a highly segregated African American school has a profound effect on a student's achievement outcomes, above and beyond the effect of individual poverty or minority status. Specifically, both the racial/ethnic and social class composition of a student's school are 1 3/4 times more important than a student's individual race/ethnicity or social class for understanding educational outcomes.*

The Equality of Educational Opportunity (EEO) study, or the "Coleman report" (Coleman et al., 1966a), is widely recognized as the most important contribution by sociologists to research on schooling (Gamoran, Secada, & Marrett, 2000). The findings from the report have shaped the sociology of education, national education policies, and wider public and scholarly opinion regarding the contributions of schools and schooling to equality and productivity in the United States. Despite past reanalyses of the Coleman data and decades of research on the effects of schools as organizations, the report's fundamental finding—that a student's family background is far more important than school social composition and school resources for understanding student outcomes—still retains much of its currency.

Indeed, this interpretation has carried over to contemporary scholars and writings, including Gamoran et al. (2000), who noted, "Though policymakers drew implications from the positive impact on learning of the proportion of White students in a school, the effect of racial composition was small compared to the great importance of individual family background" (p. 37). Similarly, with respect to school resources, Guthrie (1995) stated, "For literally decades following issuance of the report, it has been cited as evidence that added financial resources make no difference in pupil performance" (p. 260).

The availability of equal educational opportunities and the need for improved equality of educational outcomes among racial/ethnic groups were some of the main concerns of the Civil Rights Act of 1964. Section 402 of the Civil Rights Act called for a survey concerning the lack of availability of equal educational opportunity by reason of race, color, religion, or national origin in pubic educational institutions at all levels. Referred to by Mosteller and Moynihan (1972b) as one of the largest social science research projects in history, the EEO was the result of this legislation. Led by James S. Coleman, then of the Department of Social Relations at Johns Hopkins University, Ernest Q. Campbell of Vanderbilt University, and personnel from the U.S. Office of Education, the EEO was undertaken to provide empirical evidence to support and to hasten the process that had been ordered by *Brown* in 1954: to desegregate "with all deliberate speed" (Mosteller & Moynihan, 1972b).

In addition to documenting the general effect of segregation on minority students' access to equal educational opportunities and outcomes, it was assumed that the Coleman report would reveal specific inequalities between the facilities and resources available to students in predominantly minority and predominantly White schools. It was also assumed that such inequalities in educational inputs would, quite naturally, be associated with inequalities in educational outputs. However, after finding surprisingly few differences between the characteristics of schools attended by minority and White students, Coleman et al. (1966a) concluded that "schools are remarkably similar in the way they relate to the achievement of their pupils" (p. 21).

The Harvard Faculty Seminar on the Coleman Report and the resulting volume edited by Mosteller and Moynihan (1972a), *On Equality of Educational Opportunity,* along with a volume by Jencks et al. (1972), *Inequality: A Reassessment of the Effect of Family and Schooling in America,* were among the first efforts to challenge Coleman's contention that schools do not make much of a difference. Contributors to *On Equality of Educational Opportunity*, including Smith (1972) and Mosteller and Moynihan (1972b), identified important statistical miscalculations and raised noteworthy methodological criticisms, including that the sample was not properly selected, the nonresponses were too many, the number of districts and schools refusing to participate invalidated the results, the statistical techniques used were inappropriate, and the achievement tests more closely resembled tests of aptitude.

In general, though, the reanalyses and critiques that came out of the Harvard Faculty Seminar and the volumes by Mosteller and Moynihan, and Jencks and colleagues confirmed the findings from Coleman's original analyses that differences in school resources were slight and that they had only a small effect on achievement. As Jencks et al. (1972) concluded after their extensive reanalyses, "There is no evidence that school reform can substantially reduce the extent of cognitive inequality. . . . Neither school resources nor segregation has an appreciable effect on either test scores or educational attainment" (p. 8).

These reanalyses and reinterpretations further weakened and qualified the Coleman report's mild assertions regarding the potential positive effects of racial and socioeconomic integration. Bowles and Levin (1968) insisted, "We find that the conclusion that Negro achievement is positively associated with the proportion of fellow students who are white, once other influences are taken into account, is not supported by the evidence presented in the Report" (p. 23). Bowles and Levin posed some technical arguments to support this conclusion but also referred back to the original EEO report, which stated, "The effects of the student body environment upon a student's achievement appear to lie in the educational proficiency possessed by that student body, whatever its racial or ethnic composition" (Coleman et al., 1966a, p. 307). Bowles and Levin went on to state, "And in fact Coleman has emphatically stressed that the survey revealed no unique effect of racial composition on the achievement levels of nonwhites" (p. 22).

Later, scholars who study schools as organizations critiqued conceptual and technical inadequacies of Coleman's education production function models and articulated the differences between school effects—the organizational context for teaching and learning—and the effects of schooling—the experiences that students have in schools and classrooms that actually produce learning (Bidwell & Kasarda, 1980; Lee & Bryk, 1989). Most important, these researchers demonstrated that the effects of schooling are mediated by processes occurring at multiple levels of school system organization, from within-school processes, like tracking and ability grouping, to the organizational context of the school, to higher level policies imposed by district, state, and federal mandates and decisions. During the 1980s, and more prominently with the advances in multilevel modeling techniques in the 1990s, this perspective, which Gamoran et al. (2000) called the "nested layers approach," gained considerable attention as a way to understand the effects of schools and schooling.

The present study extends both of these lines of reanalysis and reconceptualization—the statistical revisionist perspective of Mosteller and Moynihan (1972a) and Jencks et al. (1972), and the nested-layers approach suggested by Bidwell and Kasarda (1980)—that have emerged in response to the Coleman report. From both a statistical and theoretical perspective, we believe that the research problems are most appropriately understood as multilevel, with a micro (within-school, or student-level) and macro (between-school, or school-level) component. The primary statistical tool that we use, the multilevel model, explicitly takes into account this hierarchical structure. We reanalyze the ninth-grade data from the EEO survey using contemporary statistical methods that were not available to Coleman, his colleagues, and past critics. This project recasts the original EEO production function models, which

Coleman and his colleagues estimated via ordinary least squares (OLS) regression, as two-level hierarchical linear models (HLMs) of the effects of (a) school-level social composition (e.g., poverty and racial/ethnic composition) and educational resources on students' verbal achievement, and (b) within-school curricular differentiation and teacher effects on the achievement gaps separating both African American and White students and students from more and less advantaged family backgrounds. The overarching question motivating this research is: Would Coleman and his colleagues have reached the same conclusions had they had available today's state-of-the-art statistical methods and theories? In particular, how might multilevel modeling techniques have changed the specification, interpretation, and conclusions of, arguably, the most important study of schools and educational equality in history?

THEORIES EXPLAINING SOCIAL CONTEXT EFFECTS AND EDUCATIONAL INEQUALITY

To what extent does the poverty and minority concentration within a school affect a student's achievement outcomes, above and beyond the effect of his or her individual poverty and minority status? Moreover, if the school's social context does matter, what are the underlying mechanisms through which it is manifested? These questions, which link the collective with the individual in educational settings, are fundamental to sociological endeavors. Though they were central questions asked by Coleman and colleagues, as Jencks and Mayer (1990) noted, they remain poorly understood.

One theory suggests that social context is linked to schools' unequal distributions of resources and opportunities. Referred to as the institutional model (Jencks & Mayer, 1990), it suggests that we may understand the potentially deleterious impacts of high-poverty and highly segregated communities by looking to the schools and other institutions serving the neighborhood. This was the primary model of inequality used in the Coleman report. Specifically, variables measuring the school organizational resources (including the overall per-pupil expenditure, the number of science laboratories, and the number of volumes per student in the school library) and the classroom-based resources (such as the teachers' years of experience, knowledge as measured by a verbal test score, and potential biases and perceptions, including the degree to which the teachers preferred teaching middle-class students) were some of the key factors used to predict differences in student outcomes across varying school contexts.

At least two other prominent models of compositional effects have been advanced by researchers. It may be the case that attending high-poverty and largely African American schools constricts students' educational opportunities through peer networks that reinforce behaviors, attitudes, and beliefs that are in opposition to traditional middle-class values regarding the importance of education. Researchers who have advanced these "epidemic" theories assume that good or bad behavior is contagious and that interactions among classmates or schoolmates are important mechanisms for shaping the academic trajectories of individuals. Beginning with the work of Wilson (1959), who explored the effects of a high school's average socioeconomic status (SES) on graduating seniors' college plans, this model has emphasized more directly the role of a student's peers in shaping educational aspirations and outcomes. Using variables that included the proportion of students within the school planning to attend college and the average number of hours that students from the school worked on homework assignments, the EEO study also measured attributes of the epidemic theory of compositional effects.

Finally, the collective socialization model holds that the social networks and relationships between adults and children within a school and neighborhood are also important resources from which students may benefit. In Coleman's later writings (1987, 1988), he argued that schools and neighborhoods with greater family resources tended to have more "social capital" to invest in the education of their children. Emphasizing both the strength of social relationships and the enforcement of norms imposed by parents and by the larger community, Coleman noted that Catholic schools with strong church and school communities provided some of the most notable examples of social capital. Of the three general theories explaining compositional effects, though, the collective socialization model was the most poorly represented within the EEO data and analyses.

As a social survey that was designed to serve as an instrument of national policy-making, the results and interpretations presented in the Coleman report offered little in the way of theory for exploring these various models underlying school context effects. As this brief discussion of the theories that have been advanced for explaining the effects of school social composition reveals, though, the EEO data and analyses provided a fairly thorough model of factors associated with the institutional and epidemic perspectives. Conceptualized and modeled from a more theoretical perspective, the EEO data could provide important insights into both the school compositional effects of race/ethnicity and social class, and the underlying models that help explain them.

RECENT METHODOLOGICAL ADVANCES IN SCHOOL EFFECTS RESEARCH

In 1966, the Coleman report was truly a state-of-the-art analysis of school effects. At that time, the OLS linear regression model, with student achievement scores regressed on variables measuring student inputs and school characteristics, was a pioneering methodology. The EEO also was one of the first major studies in the social sciences to depend on the then-infant technology of computing. The regressions that were estimated by Coleman and his colleagues were computed on the Model T of the computer industry, an IBM 1401 with 14k of memory. Because of the memory constraints of the computer, prior analyses of the Coleman data relied on only small random samples of 1,000 students in generating the results for each subgroup defined by race, region, urbanicity, and grade level.

Beyond these deficiencies related to the computer technology of the time, none of the possible criticisms seem more important than those involving the theory and analytical methods for partitioning and explaining the sources of variability in achievement attributable to student-level background characteristics and school-level characteristics. In the case of the Coleman report, it also was of fundamental interest to discover whether variability in school-level characteristics *mediated* the relationships between a student's racial/ethnic and social class background and his or her achievements in school. These problems, and many others in educational research, are related to the hierarchical or multilevel nature of the data for students and schools. The main hypotheses involve independent variables measured at the school level (such as policies, practices, and resources) and at the student level (such as background characteristics), and a dependent variable, usually achievement, measured at the student level. Until recently, statistical models that appropriately modeled the multilevel and interactive phenomena of school and classroom effects on student-level educational outcomes were not available. This had created serious methodological difficulties that had hampered the analytical and theoretical study of school effects (Raudenbush & Bryk, 1986). The analyses of the EEO data suffered from these same problems.

Although Coleman and his colleagues and others who reanalyzed the data recognized that the variance could be separated into within-school and between-school components, they had less efficient methods for partitioning the variance and had no practical methods for simultaneously modeling both levels of this hierarchical variance structure. As a result, different analysts chose different analytical paths. Coleman and his colleagues specified the student as the primary level of analysis, but others who reanalyzed the data, including Armor (1972), chose the school as the main unit of analysis. More recently, Burstein (1980) and Rogosa (1978), among others, contended that dilemmas such as this that involve the choice of which unit to analyze were addressing the wrong question and that the most appropriate and informative model would allow estimation of random variation at both levels.

Accompanying this unit-of-analysis problem are several other technical issues that compromise the EEO analyses. First, in choosing to disaggregate the higher order variables to the individual level, Coleman and his colleagues assigned the same values for each of the school measures to all students who happened to share membership within the same school. Of course, because these students share the same value on each of the school measures, Coleman and his colleagues and others who reanalyzed the data violated the assumption of independence of observations, which is requisite for all classic statistical techniques, including the OLS regression methods used in the Coleman report and other reanalyses. By using a single-level statistical model with clustered data, the estimated standard errors for the school variables were too small, leading to liberal tests of statistical significance and an inflated probability of making a Type I error. On the other hand, the approach employed by Armor (1972), who aggregated all student-level variables to the school level, threw away all the within-school variability, which represented as much as 90% of the variation in achievement for some parts of the EEO sample (Mosteller & Moynihan, 1972b).

Further, though one of the most important objectives of the EEO report was to examine how variables measured at the level of the school affected relations between achievement outcomes and student-level characteristics—most importantly, socioeconomic status and race/ethnicity—past analyses of the Coleman data did not formally explore these so-called cross-level interaction effects. Instead, Coleman and his colleagues specified separate subanalyses for the various racial/ethnic groups and regions of the country that were represented in the data set. Although these analyses did identify relations between school characteristics and achievement for each of the various racial/ethnic groups, they did not specifically document how school characteristics may have mediated, by attenuating or amplifying, the achievement gaps between minority and White students and poor and middle-class students.

These important limitations, which involved the unit-of-analysis problem and the omission of cross-level interaction effects, were solved by the emergence of multilevel, or hierarchical, models (Bryk & Raudenbush, 1992; Goldstein, 1987). Previously unavailable to the authors of the EEO and those who have reanalyzed its results, these methods have, in several respects, brought about a revolution in the analysis of school effects. Rather than choosing between the student level or school level as the primary unit of analysis, HLMs allow the researcher to simultaneously model hypotheses about effects that occur at each level. The researcher may efficiently partition the total variance into its within- and between-schools components and explain the variability that occurs at each level with appropriate measures of student and school characteristics. Taken together, these advances allow educational researchers to model more effectively how, and for whom, schools make a difference.

OBJECTIVES OF THE CURRENT STUDY

More specifically, as we demonstrate in the current study, the multilevel model may be used to reassess the key findings of Coleman et al. (1966a) and others regarding the relative effects of family background and schools. We performed this reanalysis in seven stages. These stages mirror the original steps taken by Coleman and colleagues in partitioning and explaining the variability in achievement attributable to students' backgrounds and schools. First, we began by apportioning variation between and within schools. This first stage tells us whether, and to what extent, achievement outcomes varied as a function of students and schools. Second, we tested the heterogeneity of regression assumption for the Black-White test score gap and the social class slope. That is, we assessed whether the relationship between achievement and students' social class and racial/ethnic background varied depending on which school they attended, or whether the relationship remained unchanged across schools. During this stage, we also examined the extent to which students' individual background explained between-school achievement differences.

Third, net of individual student background, we measured differences in school mean achievement outcomes associated with school-level social class and racial/ethnic composition. Fourth, we reexamined the extent to which the facilities and curriculum measures from the Coleman report might account for the overall achievement outcomes of schools. Fifth, we modeled the teacher characteristics, or classroom-based components, that may explain the school effects. Sixth, after statistically controlling student background and school and teacher resources, we modeled the student body characteristics, or peer effects.

Finally, in addition to modeling these between-school differences in school mean achievement, we attempted to explain school-to-school variability in the within-school social class and Black-White achievement differences. These cross-level interaction effects, in the theoretical tradition of the nested layers approach articulated by Gamoran et al. (2000), tested the extent to which curricular differentiation, tracking, and potential biases among teachers explained within-school inequality.[1] This modeling of within-school gaps and slopes provides a clear analytical and theoretical departure from previous analyses of the Coleman data. Rather than assuming that only school-to-school differences are related to inequality, these additional analyses assessed how teachers and schools promoted social inequality between Black and White students and lower and higher SES students attending the same school.

METHOD

DATA

We retrieved the EEO data files from the archives maintained by the Inter-University Consortium for Political and Social Research (ICPSR). The files include data from the original stratified two-stage probability sample of the public schools in the United States and the District of Columbia. Included in the files are surveys and test scores from the original EEO sample of more than 570,000 students from Grades 1, 3, 6, 9, and 12. In addition, the files include survey responses from about 4,000 principals, and survey and test results for more than 40,000 teachers. The codebook available through the ICPSR contains documentation compiled by the National Archives and Records Service, stating that it received from Johns Hopkins University only two reels of data for the teacher file and that they were labeled "1 of 4" and "2 of 4." The teacher record count for the two reels is 44,193 and the record count for the original EEO teacher file is 66,826. Therefore, it appears that the teacher data are not available in their original and complete form. Data and documentation for the district-level survey, which, most importantly, provided per-pupil expenditure information, were also missing.

For the current study, we limited our attention to the principal surveys, teacher questionnaire and test data, and student achievement and survey data for the ninth-grade cohort. The 9th- and 12th-grade data contain the widest range of information, and it was these files that were the primary focus of the original EEO analyses. We focused on 9th-grade students rather than 12th-grade students because dropping out tends to cause less overall attrition and differential selectivity across socioeconomic and racial/ethnic groups among the earlier high school grade cohorts.

The ninth-grade files from the ICPSR contained records for 134,030 students within 930 schools. Following the data-cleaning process and after restricting our attention to only those students with race/ethnicity and achievement data, we were left with a total sample of 132,065 students and 894 schools. However, primarily because of considerable missing data from the principal surveys, the sample was reduced to 50,541 students and 409 schools after including the school variables gleaned from the principal surveys. Finally, because of nonreturned surveys and the missing teacher data reels, the final sample sizes, after including both school and teacher variables, were reduced to 30,590 students and 226 schools. Though these missing data rates are high, previous analyses have been affected by significant data attrition as well. For instance, Bowles and Levin (1968) reported that only 59% of the high schools returned complete sets of the surveys, and 21% of the ninth-grade student surveys omitted information about father's education. Despite these problems of instrument and item nonresponse, examination of the general background characteristics of students and compositional and regional data for schools tabulated in Table 1 shows few differences between the total sample and the reduced samples. There were no differences on any of the student-level variables that exceeded a 10th of a standard deviation, and there was only one such school-level variable: school mean parental education. The schools in the final analytical sample were approximately 0.15 standard deviations less advantaged with respect to the aggregate measure of school mean parental education. A summary of the descrip-

Table 1. Student and School Demographic Characteristics for the Total Sample and Reduced Samples

	Total Sample			Sample After Including School Variables			Sample After Including School and Teacher Variables		
Variables	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Student-Level Characteristics									
Verbal Score	132,065	27.55	12.95	50,541	27.11	13.00	30,590	27.54	12.94
Black	132,065	0.28	0.45	50,541	0.28	0.45	30,590	0.27	0.44
White	132,065	0.58	0.49	50,541	0.59	0.49	30,590	0.61	0.40
Hispanic	132,065	0.05	0.22	50,541	0.05	0.22	30,590	0.04	0.21
American Indian	132,065	0.02	0.15	50,541	0.02	0.15	30,590	0.02	0.14
Asian American	132,065	0.01	0.11	50,541	0.01	0.09	30,590	0.00	0.06
Other Race/Ethnicity	132,065	0.06	0.23	50,541	0.05	0.23	30,590	0.05	0.22
Parental Education	113,375	0.00	1.00	43,929	-0.11	0.99	26,615	-0.09	0.98
Family Resources	131,276	0.00	1.00	50,257	-0.09	1.07	30,418	-0.05	1.05
School-Level Characteristics									
Percent Blacks	894	0.34	0.39	409	0.35	0.40	226	0.33	0.39
Family Resources	894	0.00	1.00	409	-0.12	1.00	226	-0.06	1.03
School Mean Parental Education	894	0.00	1.00	409	-0.17	0.92	226	-0.15	0.98
South	894	0.55	0.50	409	0.63	0.48	226	0.52	0.50
Metro	788	0.20	0.40	409	0.11	0.31	226	0.16	0.37

tive statistics for the student-level and school-level variables is presented in Table 2 for the analytical sample. The procedures for creating these student and school variables are described in the following section.

Table 2. Descriptive Statistics for Student- and School-Level Variables for the Analytical Sample

Variable	*n*	*M*	*SD*	Min	Max
Student-Level Variables					
Verbal Achievement Score	30,590	27.54	12.94	0	60
Number of Siblings	30,213	3.84	2.65	0	9
Two-parent Household	30,590	0.76	0.43	0	1
African American	30,590	0.27	0.44	0	1
White	30,590	0.61	0.49	0	1
Hispanic	30,590	0.04	0.21	0	1
American Indian	30,590	0.02	0.14	0	1
Asian American	30,590	0.00	0.06	0	1
Other	30,590	0.05	0.22	0	1
Parents' Education	26,615	-0.09	0.98	-2.02	2.92
Reading Material in Home	30,410	-0.19	3.19	-11.26	4.68
Family Resources	30,418	-0.05	1.05	-5.34	0.64
Urbanism of Background	30,054	-0.11	0.96	-1.26	2.89
School-Level Variables					
Social Composition					
Percent African Americans	226	0.33	0.39	0	1
School Mean Family Resources	226	-0.06	1.03	-2.98	1.34
School Mean Parental Education	226	-0.15	0.98	-2.98	4.14
Facilities and Curriculum Measures					
South	226	0.52	0.50	0	1
Metro	226	0.16	0.37	0	1
Science Laboratory Facilities	226	2.51	0.89	0	3
Extra-curricular Activities	226	11.67	3.04	2	18
Comprehensiveness of the Curriculum	226	3.83	1.45	0	6
Guidance Counselors	226	2.35	2.27	0	9
Accelerated Curriculum	226	0.76	0.96	0	3
School Enrollment	226	6.46	0.77	2.2	8.01
Tracking	226	0.58	0.49	0	1
Volumes per Student	226	6.82	5.27	0.1	35.71
Movement between Tracks	226	0.05	0.07	0	0.47
Expenditure (Average Teacher Salary)	226	8.63	0.20	8.16	9.1
Teacher Variables					
Teachers' Average Years of Experience	226	3.26	0.54	1.61	5.07
Preference for Middle Class	226	-0.14	0.84	-3.13	1.87
Localism	226	-0.33	1.17	-3.3	2.21
Percent White Teachers	226	0.65	0.45	0	1
Teacher Education Level	226	0.01	0.42	-2.06	1.29
Family Education Level	226	-0.02	0.38	-1.38	1.23
Verbal Score	226	-0.13	0.54	-1.69	0.64
Student Body Characteristics					
Proportion of Families that Own Encyclopedia	226	0.66	0.21	0.14	0.98
Transfers	226	-0.19	0.30	-0.81	0.78
Students Planning to Attend College	226	-0.09	0.34	-2.13	0.74
Hours of Homework	226	0.04	0.28	-0.65	0.91
Attendance	226	0.02	0.31	-1.44	0.71

MEASURES

We developed student-level and school-level variables using the same methodology employed in the original EEO report. Our procedures relied on the same data elements and methods as described on pages iii–vii of the supplemental appendix to the EEO (Coleman et al., 1966b). The analytical procedure for developing composite variables also replicated the approach used by Coleman and colleagues: We first standardized each item and then combined them to form the composite scores. The measures that we used in our models were those that formed the core of the EEO analyses of the relation of school, teacher, and student body characteristics to achievement, which were presented in chapter 3 of the original report.

Student-level variables. To study the relationships between student background factors and achievement, Coleman and his colleagues defined the following eight variables based on student-reported information from the surveys: (1) urbanism of background, (2) parents' education, (3) structural integrity of the home, (4) smallness of family, (5) items in the home, (6) reading materials in the home, (7) parents' interest, and (8) parents' educational desires. Coleman and colleagues referred to the first six variables as "objective background family factors" and the last two variables as "subjective background family factors." The major focus of the EEO analyses of school effects, and the focus of the present study was the first six variables: the objective background family factors.[2]

The only single indicator in this set of variables was the smallness of family variable, which was based on the student's number of siblings. The remaining variables were linear composites of several items from the student questionnaire. Urbanism of background was based on two items from the student questionnaire related to the community in which the student and mother grew up. Parents' education level was created by taking the average of the mother's and father's reported education level. The structural integrity of the home was based on two questions concerning whether the mother and father resided with the student at home. This variable was further dichotomized, assigning a value of 1 if the student came from a two-parent household, 0 otherwise.

The student's number of siblings was based on one item from the student questionnaire and had values ranging from 0 siblings to 9 or more siblings. The items in the home variable defined the family resources available within the home. This composite variable was based on whether the student's family owned the following: television, telephone, record player, refrigerator, automobile, and vacuum cleaner. The last objective background variable, reading material in the home, was created by combining the students' reports of whether the family possessed the following five items: dictionary, encyclopedia, daily newspaper, magazines, and books. Five additional student-level variables concerning ethnic background were developed for the analysis: African American, Hispanic, American Indian, Asian American, White, and other race.[3] These variables were coded as 1 if the student belonged to that ethnic category, 0 if not.

School-level variables. Twenty-seven school-level variables, which are summarized in Table 1, were included in the study. Three variables described the social composition of the schools: percent African American, school mean family resources, and school mean parental education.[4] The percent African American ranged from 0% to 100%, with a grand mean of 33%. The school mean family resources was computed by obtaining the average of the student-level family resources scores for each school and further standardizing the obtained aggregate school-level means. School mean parental education was the average of the student-level parents' education variable.

Twelve of the school-level variables were defined by Coleman and colleagues as the "facilities and curriculum measures," and 11 were taken from the principal's questionnaire. Also included as a facilities and curriculum measure was per-pupil expenditure. Because the district data on per-pupil expenditure were not available, we used the school-level average of the teacher salaries as an estimate of each school's per-pupil expenditure.[5] The final school-level average of the teacher salaries was transformed by taking its natural log.

Another seven variables were obtained from the teacher questionnaire and represented school-level averages of teacher characteristics. Finally, the last six school variables in Table 1, referred to by Coleman et al. (1966a) as the "student body characteristics," were obtained from the student questionnaire and also represented school-level average measures.

The 12 facilities and curriculum measures included the following six single indicator variables: (1) geographic region, coded as 1 for South and 0 for North[6]; (2) number of college guidance counselors at the high school, coded from 0 to 9 in the analytical sample; (3) the availability of an accelerated curriculum, coded as 0 for no accelerated curriculum, 1 for an accelerated curriculum in one or two subjects, 2 for an accelerated curriculum in several subjects, and 3 for an accelerated curriculum available in all subjects; (4) the overall school enrollment, which was transformed by taking its natural log; (5) the presence of some form of tracking in the school, coded as 1 when tracking was used and 0 when it was not used; and (6) the school location, coded as 1 for a metropolitan location and 0 for a nonmetropolitan location.[7]

The remaining facilities and curriculum measures were composite scores, previously referred to within the EEO report as the "special measures." The composite scores, or special measures, are described next.

Science laboratory facilities was based on the combination of the following three types of science laboratories in the school: biology, chemistry, and physics. The item was coded as 0 for no laboratory facilities, 1 if the school had any one of the three laboratory facilities, 2 for two of the three laboratory facilities, and 3 if the school had all three types of science laboratory facilities.

Extracurricular activities was based on the number of extracurricular activities available at the school. In the analytical sample, this item ranged from 2 to 18. The extracurricular activities included student government, newspaper, annual, boys' interschool athletics, girls' interschool athletics, boys' intramural athletics, girls' intramural athletics, band, chorus, honor society, subject clubs, chess clubs, hobby clubs, drama, debate, social dances, military cadets, service clubs, and religious clubs.

Comprehensiveness of the curriculum was determined by the number of alternative curricular tracks available at the school: (1) college preparatory, (2) commercial, (3) general, (4) vocational, (5) agriculture, and (6) industrial arts. The variable ranged from 0, or no alternate tracks, to 6 in the analytical sample.

Volumes per student was a composite obtained by dividing the number of volumes within the school library by the total number of students enrolled in the school. The grand mean for our analytical sample was nearly 7 volumes per student.

Movement between tracks was derived from the combination of two items from the principal's questionnaire that asked for the percentage of students who moved from one academic track to a higher track since September 1964, and the percentage of students who moved from one academic track to a lower track since September 1964. The composite score ranged from 0% to 47% track movement, with an overall mean of 5%.

Seven variables, referred to by Coleman and colleagues as the "teacher characteristics," were obtained from the teacher questionnaire. Akin to the approach taken in the EEO report, these variables were coded individually for each teacher in the sample and then aggregated by school to create school-level means for each of the measures. The following four were formed based on school-level averages of continuous single-indicator teacher variables: (1) average years of experience, which was transformed to the square root of the number of years teaching reported by the teacher, (2) proportion of White teachers (coded as 1 for White and 0 for the categories African American, Asian American, and other race),

(3) teacher verbal score, and (4) teacher education level with five categories ranging from no degree (coded as 0) to doctoral degree (coded as 5). The years of experience and proportion White teachers were based on simple school-level aggregates of the teacher data. Teacher verbal score and teacher education level were standardized before obtaining the school average measure.

The remaining three variables were linear composites of two or more standardized items. In all cases, the composites were computed by standardizing to a mean of 0 and standard deviation of 1 the teachers' responses to each item forming the composite. After standardizing, we then computed a teacher-specific mean of the items forming the composite. Finally, we aggregated the teacher data by school and computed an aggregate mean as the final school-level measure.

Localism was obtained by aggregating teachers' responses to the following three items in the questionnaire: Where have you spent most of your life? Where did you graduate from high school? What is the location of the undergraduate college institution attended? The survey response options ranged from a location that was within the locale in which the teacher currently taught, to a location out of the country. Locations closer to the teacher's current school were coded as higher, and locations farther away were coded lower. In addition, the number of years of full-time teaching experience at the current school was divided by the total number of years of full-time teaching experience to obtain the proportion of total years teaching spent in the current locale.

Preference for middle-class students was a composite of three variables. The first variable asked teachers about the type of high school they preferred to work in, with the following choices and coding: (1) a commercial or business school; (2) a vocational, technical, or trade school; (3) a special curriculum school designed to serve the culturally disadvantaged; (4) a comprehensive school; (5) an academic school with strong emphasis on college preparation. The second variable asked teachers about the preferred choice of school settings, with the following seven alternatives and codes: (1) children of rural families; (2) all children of factory and other blue-collar workers; (3) mostly children of factory and other blue-collar workers; (4) children from a general cross-section of the community; (5), mostly children of professional and white-collar workers; (6) all children of professional and white-collar workers. Finally, the third variable provided four choices and asked teachers about their preferred student ability level to teach or counsel. The four choices and corresponding codes were: (1) a low-ability group; (2) a mixed-ability group; (3) an average-ability group; (4) a high-ability group.

Family education level was obtained by taking the average education level of the teacher's mother and father.

Finally, the variables referred to by Coleman et al. (1966a) as the "student body characteristics" were all student-level single-indicator variables averaged within schools to create school-level aggregates. The following five student body characteristics were included: (1) proportion of families who own an encyclopedia, (2) transfers, (3) students planning to attend college, (4) hours spent on homework, and (5) attendance. The variable "families who own an encyclopedia" was a simple school-level aggregate of the student-level dummy code indicating that the family owned an encyclopedia. The remaining four student variables, which provided a range of four to seven alternative response options, were standardized before taking the average per school.

Transfer consisted of the following five response options indicating the frequency with which the student had transferred schools: (1) never, (2) once, (3) twice, (4) three times, (5) four times or more.

Students planning to attend college had the following four alternatives and codes: (1) definitely not, (2) probably not, (3) probably yes, (4) definitely yes.

Average hours of homework consisted of seven options and codes: (7) 4 or more hours a day, (6) about 3 hours a day, (5) about 2 hours a day, (4) about 1.5 hours a day, (3) about 1 hour a day, (2) about a half hour a day, (1) none or almost none.

Attendance consisted of the following alternatives and codes, with higher values indicating more absences and greater attendance problems: (5) 16 or more days, (4) 7–15 days, (3) 3–6 days, (2) 1–2 days, (1) none.[8]

DEPENDENT VARIABLE

The dependent variable in the multilevel analysis was a measure of student achievement. Consistent with the original Coleman report, the criterion of achievement that we used was the student's score on a standardized verbal ability test—that is, a vocabulary test measuring verbal skills. In the current analysis, we refer to this score as verbal achievement. The scores on the verbal achievement outcome ranged from 0 to a maximum of 60, with a *SD* of 12.94.

PROCEDURE

Rather than estimating separate analytical models for African American and White students, as in the original report and prior reanalyses, we estimated an overall model incorporating the entire student and school samples. These models allowed us to measure the consequences of within- and between-school differences for the academic outcomes of all students from the national sample within an integrated model.[9] In addition, as we explain, this integrated model allowed us to compare the magnitudes of the school-level and student-level coefficients for race/ethnicity, specifically those for African Americans. This facilitated a direct assessment of compositional effects. To demonstrate the differences and similarities between the HLM and OLS approaches, we ran the same series of models using the two methods.[10]

We began by specifying an unconditional multilevel model, with no student or school predictors of the verbal achievement outcome. This model decomposes the variance in the outcome into its within- and between-school components and provides an estimate of the proportion of variability in verbal achievement that can be explained by differences across schools. The second set of multilevel models that we estimated introduced the student objective background characteristics as predictors of achievement. These models estimated inequalities in students' outcomes and accounted for within- and between-school variability associated with their individual and family backgrounds.

After assessing the variability in achievement associated with student-level characteristics, the third model turned to the compositional effects of the percent African Americans attending the school, and the school mean family resources and parent education levels of the students' families. These models estimated the compositional effects of these characteristics net of family background. Fourth, we modeled the facilities and school curriculum measures, the primary school inputs from the Coleman report, as predictors of school-to-school differences in education production. In this model, a subset of the facilities and school curriculum variables—tracking, movement between tracks, and comprehensiveness of the curriculum—was also entered to explain school-to-school differences in the within-school Black-White achievement gap and the within-school relationship between students' family resources and achievement.

The next cluster of school-level variables that we accounted for in the fifth model was teacher characteristics. In this model, we also used the teacher variable *preference for middle-class students* as a predictor of variability in the within-school Black-White test score gap and the family resources-achievement slope. Finally, in the sixth model, we entered the school-level student body characteristics. In this final comprehensive model, we evaluated the effects of school social composition and the extent to which they may be explained by the variables representing the school curriculum and facilities, the schools' teaching staffs, and the schools' student body characteristics.

The multilevel models allow for decomposition of the person-level and school-level effects of social class and race/ethnicity into separate levels (student and school compositional) and components. Within the multilevel framework, the compositional effect can be defined as the extent to which the magnitude of the organizational-level relationship, β_b, differs from the person-level relationship, β_w (Raudenbush & Bryk, 2002). The compositional effect can thus be given as $\beta_c = \beta_b - \beta_w$.

The compositional effect may be estimated in two distinct ways, which differ based on how one chooses to center the level 1 student variable. In both cases, the person level X_{ij} is included in the level 1 model, and its aggregate, the school-level mean of the student X_{ij}s, is included in the level 2 model as a predictor of the school mean achievement intercept. When one chooses group-mean centering, the level 1 student social class or race/ethnicity variable is centered on its corresponding level 2 school social class or race/ethnicity mean, and the intercept can be interpreted as the unadjusted mean for school j. When grand-mean centering is selected, the student variable is centered on the school-level grand mean and, akin to the classical analysis of covariance model, the intercept is interpreted as an adjusted mean for school j. In the former case, the relationship between X_{ij} and Y_{ij} is directly decomposed into its within-, β_w, and between-group, β_b, components, and the compositional effect can be derived by simple subtraction, $\beta_c = \beta_b - \beta_w$. In the latter case, the compositional effect is estimated directly, and β_b is obtained by addition, $\beta_b = \beta_c + \beta_w$.

Consistent with prior research summarized by Jencks and Mayer (1990), we hypothesized that β_b and β_w would be of comparable magnitudes. For most student-level variables, we elected to use grand-mean centering. However, for those student-level predictors that we modeled as randomly varying across schools, we chose group-mean centering. We adopted the group-mean centering approach when estimating the variance, of the level-one coefficients, because we assumed that the group means of the various predictors, *X*, would vary systematically across schools. In general, if the means of the *X*s vary systematically across level-two units, the choice of centering (i.e., group-mean centering vs. centering on a constant) will make a difference in estimating, and Raudenbush and Bryk (2002) recommended group-mean centering to detect and estimate properly the slope heterogeneity.[11]

RESULTS

Our preliminary analyses contrasted the original estimates by Coleman et al. (1966a) of the proportion of variability in achievement that was within and between schools with our contemporary estimates. Coleman et al. (1966a) had originally calculated the percent of total variance in individual verbal achievement that lay between schools in Table 3.2A.1 on page 326 of the EEO (Coleman et al., 1966a). Representing the total variation between students as SS_T, it can be partitioned as $SS_T = SS_B + SS_W$, where SS_B is the sum of squared deviations of school mean achievement from the overall mean, and SS_W is the sum of squared deviations of student scores within a school from the school mean.

In an analysis of variance (ANOVA), the ratio of the between-schools sum of squares (SS_B) relative to the total sum of squares (SS_T) is equivalent to the correlation ratio η^2,

$$\eta^2 = \frac{\sum_j n_j (\bar{x}_{.j} - \bar{x}_{..})^2}{\sum_{i,j} (x_{ij} - \bar{x}_{..})^2} .$$

The proportion of variation in verbal achievement that lies between schools, SS_B, was expressed by Coleman and his colleagues as a percentage of SS_T (that is,) for each of the racial/ethnic groups and regions across all grades. We report these figures in the first column of Table 3. In addition, we provide equivalent estimates that are based on the total Grade 9 sample that we extracted from the data files. The figures are roughly equivalent, suggesting that the data we extracted and the data from the original EEO sample do not appear to yield important differences. The one key difference, though, is that our current estimates also provide an indication of the overall—across all racial/ethnic groups and regions within the national sample—percent of between-school variance. This national estimate from the Coleman data of over 33% is notably larger than the previous percentages—between approximately 8.5% and 18%—that were reported by Coleman for the various subsamples of racial/ethnic groups and regions.

The intraclass correlation coefficients (ICCs) obtained in HLM for the unconditional model provide estimates of the between-school achievement variability and are also reported in Table 5 for both the complete and the final analytical Grade 9 samples. When we compare the outcomes that are based on the complete sample and derived from the ANOVA and HLM estimates, the results from the HLM analysis indicate somewhat larger percentages of variation in the achievement outcome that are attributable to schools. This is particularly the case for the sample of African American students in both the North and South. Across all racial/ethnic groups and regions, the restricted maximum likelihood estimates for the percent of between-school variation found in the complete sample are both approximately 36%. These estimates derived from HLM are somewhat larger than the figure we derived from the ANOVA-based analysis, which suggested that approximately 33% of the variation lay between schools.

Finally, the restricted maximum likelihood results for the final analytical sample of 30,590 students in 226 schools, which we analyzed in our HLM and OLS models reported in Tables 4 and 5, revealed slightly higher percentages of approximately 40% between-school variation relative to the outcome of about 36% reported for the complete sample of 132,065 students in 894 schools. Though there is somewhat more between-school variation within our analytical subsample than within the complete data, the results suggest that both the analysis of the full national data set across all racial/ethnic groups and regions, and the application restricted maximum likelihood estimation via HLM contribute to our finding that a considerably higher overall proportion of variance—as much as 40%—is attributable to differences across schools.

Table 3. Comparison of the Percent of Total Variance in Verbal Achievement that Lies Between Schools for the Grade 9 Sample

	Percent Between-School Variance Derived from ANOVA Estimates		Percent Between-School Variance Derived from Restricted Maximum Likelihood HLM Estimates	
Strata	Current Coleman et al. (1966, p 326) (Table 3.2A.1)	Current Estimates for the Complete Data[a]	Current Estimates for the Complete Data[a]	Current Estimates for the Analytical Sample[b]
Whites North	8.51	12.17	13.80	13.01
Whites South	9.12	10.60	11.56	10.27
Blacks North	13.37	12.61	18.21	21.79
Blacks South	17.98	18.37	22.54	27.42
Total	N/A	33.40	36.15	39.83

Note. [a] The complete data set is composed of 132,065 students in 894 schools. See descriptive statistics for this sample in Table 1 in columns labeled "Total Sample."
[b]The analytical sample is composed of 30,590 students in 226 schools. See descriptive statistics for this sample in Table 1 in columns labeled "Sample After Including School and Teacher Variables."

THE EXPLANATORY MULTILEVEL AND OLS REGRESSION MODELS

Table 4 displays the maximum likelihood results from the multilevel analyses, starting with the null, or unconditional, model to Model 6. The first analytical model, the null multilevel model with no student- or school-level predictors, shows the overall average value on the outcome measure, partitions the variance in the outcome into its between- and within-school components, and tests whether there is a statistically significant amount of between-school variance to model with independent variables. In comparison, the null model for our OLS regression analyses in Table 5 yielded an average verbal achievement score of 27.55. This model does not explicitly partition the school- and student-level variance into separate components. The analysis is specified at the level of the student, and the OLS models that follow include both student and school variables as predictors of differences among students in the verbal achievement outcome.

For the verbal achievement outcome, the unconditional multilevel model summarized in Table 4 yielded an average score of 25.17. The model also revealed that there was a statistically significant, $\chi^2(225, N = 226) = 19{,}549.86$, amount of level 2 variability potentially explainable by school-level characteristics. Thus, we began the specification of our multilevel school-level explanatory models and our OLS regression models.

Model 1: Objective family background variables as predictors. Our next steps involved attempting to control the objective family background factors and using the school-level compositional variables, facilities and curriculum measures, teacher characteristics, and student body characteristics as predictors of verbal achievement. With the exception of the African American and family resources predictors, our HLM models treated the student-level race/ethnicity indicators and objective family background variables as fixed slopes. That is, it was assumed that the effect of most student-level predictors was homogeneous across schools. We chose this model because of both practical and theoretical considerations. From a practical standpoint, there were a number of schools that did not serve students of Hispanic, American Indian, or Asian American backgrounds. Having no variability on these student-level indicators, in many cases, it was not possible to model these race/ethnicity indicators as sources of random variation within schools. In addition, like the original EEO, from an analytical and theoretical perspective, the primary focus of the current study was on the sources of between-school differences in mean verbal achievement rather than on processes of within-school achievement differentiation. The two exceptions were, of course, the within-school inequalities associated with social class, as measured by family resources, and a student's status as an African American.[12]

In using a grand-mean centering transformation of the student-level variables included in our multilevel models, we generated a school-level mean achievement intercept that can be interpreted as a statistically "adjusted" mean for school j. That is, after adjusting β_0 for differences in the schools' distributions of each student-level predictor, we can estimate the value-added effects of the school-level predictors net of student background. Specifying a random-intercept model, we used the school compositional variables, facilities and curriculum measures, teacher characteristics, and student body characteristics as predictors of between-school mean verbal achievement differences. This model specification most clearly helped us answer the question, Does the social class composition and concentration of African American students within a school affect a student's achievement outcomes, above and beyond the effect of his or her individual social class and minority status?

In the initial prediction model, Model 1, the objective family background characteristics explained 68.33% of the between-school variance. Therefore, the student-level predictors did account for considerable between-school variability, but a statistically significant amount of between-school variability remained even after controlling for all the measures of family background.

All the family background measures were statistically significant predictors of verbal achievement. On average, African American students obtained test scores that were 5.49 points, or 0.42 standard deviations (*SD*s), lower than White students, after controlling for other family characteristics. Those with higher verbal achievement scores tended to be White students from families with higher levels of parental education, fewer siblings, greater family resources, more literacy-rich home environments, and both parents residing at home. Finally, the variable urbanism of background was a positive predictor of achievement: The more urban the community in which the student and mother grew up, the higher the score on the verbal test.

The HLM results also showed that there were statistically significant, $\chi^2(175, N = 176) = 304.41$, level 2 differences across schools in the Black-White achievement gap and in the family resources slope, $\chi^2(175, N = 176) = 254.08$. These results provided evidence that the social distribution of achievement varied across schools. That is, some schools were more equitable and some were less equitable with respect to both race and social class.

The OLS regression model in Table 5 reveals similar results, in that the magnitudes of most coefficients for the student background characteristics are similar to those in the HLM model. The estimates from this model and those from the HLM differ in three notable ways, though. First, the standard errors for the coefficients in the OLS regression are quite a bit smaller than those from the multilevel model. The OLS model assumes that observations across students are independent and have a common variance. This assumption, though, is not likely to hold because the students in the EEO data set do not represent a simple random sample but are instead clustered within schools. As a result of this clustering, the students are more alike than they would be from a simple random sample. Because the OLS model assumes an independent error structure that does not exist, the within-school homogeneity among students creates the illusion of greater reliability and stability of the coefficient estimates, which results in underestimates of the standard errors and associated tests of statistical significance that are too liberal.

Second, of primary interest in the HLM is the variance explained among schools accounted for by the predictors. After partitioning the variance into its within- and between-school components, the multilevel model also describes directly how much variance was observed at the school and the student level. The OLS model does not make such distinctions. Instead the R^2 for the OLS model, 37.14%, refers simply to the overall variance explained in the outcome by the predictors.

Finally, the OLS model assumes homogeneity of regression, but the results from the previous HLM showed that this assumption does not hold. The relationships between achievement and both the student-level African American indicator and family resources measure vary across schools. HLM enabled us to estimate a separate set of regression coefficients for each school, and then, as we demonstrate in some of the models that follow, to model variation across schools in their sets of coefficients as multivariate outcomes that may be explained by school-level features.

Model 2: Adding school social composition predictors. After having found from the unconditional HLM that the mean verbal achievement outcome differed across schools and, from Model 1 in Table 4, that a statistically significant amount of between-school variability remained to be explained above and beyond that accounted for by student-level background, the next step was to model this remaining variability using school-level predictors of achievement. We began by including the compositional variables in Model 2.

The magnitudes of the coefficients for the compositional effects of percent Blacks and school mean family resources were considerable. The multilevel model shows the school contextual effects on verbal achievement controlling for the collection of student background characteristics. With group-mean centering, the compositional effects for percent

Blacks and school mean family resources can be derived by simple subtraction, $\beta_c = \beta_b - \beta_w$, or -5.38 = -9.66 – (-4.28). After controlling for student background characteristics, the between-school effect of percent African American, -9.66, was substantial in magnitude and was statistically significant. Indeed, this model suggested that the achievement difference between a school with no African American students and a school of 100% African American enrollment was 1 1/4 times greater than the achievement difference between an African American student and a non–African American student.

The compositional effect of school mean family resources, 1.57, was more than 3 times that of the student-level effect of family resources. There was no compositional effect for school mean parental education because the individual effect of parental education was greater in magnitude than the school-level effect for mean parental education. The inclusion of these student compositional effects explained 92% of the between-school variance in the verbal achievement outcomes, a 50-percentage-point increase beyond that explained by individual student background in Model 1.

The OLS model can also be used to estimate compositional effects. In general, the OLS estimates are unbiased but not as efficient as the HLM estimators (Raudenbush & Bryk, 2002). By subtracting the student-level coefficient of -5.94 for the group-mean centered African American status dummy code from the coefficient of -10.21 for the school-level percent Blacks predictor, the OLS model estimate for the compositional effect is -4.27. Though the coefficient for percent African American is nonzero, because it is smaller than the individual student-level effect of being Black, no compositional effect is present. Similarly, there was no compositional effect for mean parental education. The compositional effect of school mean family resources was 3 times the magnitude of the student-level effect. Modeling the same student and school predictors as those used in the HLM, the OLS model explained 39.17% of the variance in the outcome, or an added 2% of the variability in verbal achievement.

Understanding why the proportion of variance accounted for by school composition is so different between the HLM and OLS models requires careful consideration of how the variance in the verbal achievement outcome was partitioned in each instance. In the case of the multilevel random-intercept model, the compositional variables modeled at level 2 only account for parameter variation, τ_{00}, among the true school means, β_{0j} (Raudenbush & Bryk, 2002). The 92% variance explained by Model 2, which added the school composition measures, suggests that even after adjusting for the student background characteristics, a large proportion of the variation among true school means is related to differences in the social contexts of schools. In comparison, the variance-explained statistic for OLS uses as a denominator the total variability in the verbal achievement outcome, including both within-school and between-school variation. As Raudenbush and Bryk (2002) noted, the within-school variation reflects individual effects and errors of measurement in the outcome, both of which are unexplainable by school compositional features. Thus, when judged against this standard, the 2% of additional variance explained by school composition in the OLS model appears deceptively small.

Model 3: Adding school facilities and curriculum predictors. Compared with the previous model, the HLM labeled Model 3 in Table 4, which added the special measures and indicators of school facilities and curriculum, accounted for less than 1% of additional between-school variance in school mean achievement. The variables did, however, explain a portion of the social compositional effects; the magnitudes of the school percent Black, the school mean family resources, and mean parental education coefficients decreased relative to the previous model. After adjusting for the student background characteristics and controlling for school social composition, only one variable, the indicator of the school location in the South, was a statistically significant level 2 predictor of verbal achievement. Attending a school in the South was associated with a deficit of approximately 2.4 points in verbal achievement.

The other key outcomes of Model 3 are for the school-level prediction models for the family resources and Black slopes. In both school-level models, we employed the facilities and curriculum measures that we hypothesized were associated with potential within-school inequalities related to social class and race/ethnicity. These included the measures related to tracking, namely the tracking and movement between tracks variables, and curricular differentiation as measured by the comprehensiveness of the curriculum variable. Curricular differentiation and tracking did not account for school-to-school differences in the Black-White achievement gap, but curricular differentiation did explain differences among schools in their family resources slopes. As indicated by the statistically significant coefficient of 0.16 for the comprehensiveness of the curriculum measure predicting the family resources slope, schools with a broader array of curricular track offerings had steeper family resources achievement slopes. That is, schools that had greater curricular differentiation tended to exacerbate inequalities in achievement related to student social class.

The school facilities and curriculum measures entered as predictors in the OLS regression Model 3 explained nearly 2% of additional variation in verbal achievement beyond the previous OLS model. According to

Table 4. Hierarchical Linear Models Predicting Verbal Achievement

	Null Model			Model 1			Model 2		
	Effect	*SE*	*t*	Effect	*SE*	*t*	Effect	*SE*	*t*
Effects of Within-School Variables									
Verbal achievement score	***25.17	0.56	44.7	***26.72	0.46	58.71	***26.60	0.19	139.50
Number of siblings				***-0.49	0.03	-16.75	***-0.47	0.03	-15.88
Two-parent household				***1.49	0.17	9.08	***1.43	0.17	8.66
Black				***-4.57	0.38	-11.86	***-4.28	0.39	-10.90
Hispanic				***-4.55	0.39	-11.61	***-4.69	0.39	-12.12
American Indian				***-4.79	0.57	-8.44	***-4.33	0.58	-7.41
Asian American				***-4.72	1.40	-3.38	***-4.70	1.37	-3.42
Other				***-5.54	0.37	-15.16	***-5.71	0.36	-15.74
Parents' education				***2.19	0.10	21.29	***2.17	0.10	20.92
Reading material in home				***0.34	0.03	11.07	***0.33	0.03	10.72
Family resources				***0.44	0.10	4.52	***0.47	0.10	4.91
Urbanism of background				***0.71	0.12	5.86	***0.69	0.12	5.84
Effects of Between-School Variables									
School Mean Achievement									
Percent Blacks							***-9.66	1.00	-9.61
School mean family resources							***1.93	0.46	4.15
School mean parental education							***1.06	0.25	4.31
South									
Metro									
Science laboratory facilities									
Extracurricular activities									
Comprehensiveness of the curriculum									
Guidance counselors									
Accelerated curriculum									
School enrollment									
Tracking									

Volumes per student									
Movement between tracks									
Expenditure (average teacher salary)									
Average years of experience									
Preference for middle class									
Localism									
Percent White teachers									
Teacher education level									
Family education level									
Verbal score									
Prop families own encyclopedia									
Transfers									
Proportion planning to attend college									
Average hours homework									
Attendance									
Black Slope									
Comprehensiveness of the curriculum									
Tracking									
Movement between tracks									
Preference for middle class									
Family Resources Slope									
Comprehensiveness of the curriculum									
Tracking									
Movement between tracks									
Preference for middle class									
Variation Between Schools	Estimate	X^2	*df*	Estimate	X^2	*df*	Estimate	X^2	*df*
Intercept	70.20	***19549.86	225	40.71	***8613.74	167	5.80	***1173.41	164
Black slope				5.83	*** 266.43	167	5.91	***272.95	167
Family resources slope				0.35	***209.14	167	0.40	**209.83	167
Variation Within Schools	106.05			93.13			93.12		
Between-schools: R^2				42.00			92.00		

(Table 4 continued)

	Model 3			Model 4			Model 5		
	Effect	*SE*	*t*	Effect	*SE*	*t*	Effect	*SE*	*t*
Effects of Within-School Variables									
Verbal achievement score	***26.60	0.18	146.76	***26.60	0.18	149.43	***26.63	0.17	156.05
Number of siblings	***-0.47	0.03	-15.94	***-0.47	0.03	-15.78	***-0.47	0.03	-15.76
Two-parent household	***1.44	0.17	8.71	***1.43	0.17	8.61	***1.42	0.17	8.57
Black	***-4.31	0.40	-10.70	***-4.31	0.36	-11.91	***-4.28	0.36	-11.74
Hispanic	***-4.70	0.39	-12.17	***-4.60	0.38	-11.99	***-4.60	0.38	-12.01
American Indian	***-4.60	0.57	-8.13	***-4.59	0.56	-8.17	***-4.69	0.57	-8.26
Asian American	***-4.73	1.38	-3.42	***-4.66	1.38	-3.38	***-4.66	1.38	-3.38
Other	***-5.75	0.36	-15.75	***-5.43	0.37	-14.63	***-5.42	0.37	-14.59
Parents' education	***2.18	0.10	20.93	***2.17	0.10	20.74	***2.17	0.10	20.78
Reading material in home	***0.32	0.03	10.66	***0.32	0.03	10.51	***0.32	0.03	10.44
Family resources	***0.45	0.10	4.63	***0.46	0.10	4.54	***0.47	0.10	4.60
Urbanism of background	***0.69	0.12	5.71	***0.70	0.12	5.79	***0.71	0.12	5.88
Effects of Between-School Variables									
School Mean Achievement									
Percent Blacks	***-9.39	0.76	-12.33	***-11.77	1.47	-7.99	***-11.79	1.44	-8.20
School mean family resources	***1.53	0.39	3.96	***1.65	0.39	4.19	**1.35	0.46	2.93
School mean parental education	***0.83	0.24	3.49	0.43	0.27	1.63	-0.30	0.32	-0.96
South	***-2.39	0.53	-4.48	**-2.11	0.58	-3.62	***-2.55	0.56	-4.52
Metro	-0.57	0.51	-1.13	-0.30	0.56	-0.53	0.03	0.49	0.06
Science laboratory facilities	0.40	0.31	1.28	0.42	0.32	1.33	0.31	0.29	1.05
Extracurricular activities	0.05	0.08	0.60	0.01	0.09	0.07	0.01	0.08	0.19
Comprehensiveness of the curriculum	-0.26	0.17	-1.51	-0.34	0.18	-1.83	-0.26	0.17	-1.54
Guidance counselors	-0.06	0.13	-0.49	-0.08	0.13	-0.64	-0.06	0.12	-0.47
Accelerated curriculum	0.05	0.20	0.27	0.16	0.20	0.79	0.04	0.20	0.21
School enrollment	0.08	0.39	0.22	0.28	0.38	0.74	0.21	0.36	0.59
Tracking	0.65	0.52	1.25	0.40	0.52	0.77	0.12	0.47	0.25
Volumes per student	0.07	0.04	1.73	0.06	0.04	1.54	0.03	0.04	0.86

Movement between tracks	-5.31	3.36	-1.58	-3.77	3.31	-1.14	-2.76	3.05	-0.90
Expenditure (average teacher salary)	-1.14	1.63	-0.70	0.16	1.79	0.09	1.89	1.74	1.09
Average years of experience				0.39	0.33	1.18	0.38	0.35	1.07
Preference for middle class				0.50	0.35	1.41	0.44	0.33	1.34
Localism				-0.30	0.15	-1.95	**-0.48	0.14	-3.35
Percent White teachers				*-3.55	1.55	-2.29	-1.46	1.58	-0.93
Teacher education level				-0.86	0.56	-1.54	*-1.13	0.53	-2.13
Family education level				0.46	0.58	0.79	0.46	0.57	0.80
Verbal score				1.08	0.71	1.52	1.08	0.62	1.75
Proportion families own encyclopedia							0.71	2.25	0.32
Transfers							-1.29	0.74	-1.73
Proportion planning to attend college							***3.15	0.70	4.51
Average hours homework							0.66	0.72	0.92
Attendance							0.14	0.60	0.23
Black Slope									
Comprehensiveness of the curriculum	0.01	0.24	0.03	0.22	0.24	0.92	0.21	0.24	0.91
Tracking	1.01	0.90	1.13	1.54	0.81	1.90	1.48	0.81	1.83
Movement between tracks	-6.62	6.36	-1.04	-6.55	5.73	-1.14	-6.31	5.69	-1.11
Preference for middle class				***-2.21	0.36	-6.08	***-2.19	0.37	-5.98
Family Resources Slope									
Comprehensiveness of the curriculum	**0.16	0.06	2.80	*0.13	0.06	2.08	*0.13	0.06	2.09
Tracking	-0.28	0.23	-1.22	-0.36	0.24	-1.54	-0.37	0.24	-1.56
Movement between tracks	2.08	1.47	1.41	2.13	1.50	1.42	2.16	1.48	1.46
Preference for middle class	*0.27	0.12	2.27	*0.29	0.12	2.33			
Variation Between Schools	Estimate	χ^2	*df*	Estimate	χ^2	*df*	Estimate	χ^2	*df*
School mean achievement	5.35	***1133.64	152	5.22	***1097.19	145	4.59	***984.71	140
Black slope 5.82	***263.23	164	3.35	**210.34	163	3.42	**240.26	163	
Family resources slope	0.36	*197.27	164	0.31	189.39	163	0.30	189.28	163
Variation Within Schools	93.10			93.10			93.10		
Between-Schools: R^2	92.62			92.80			93.70		

*$p < .05$. **$p < .01$. *** $p < .001$.

Table 5. Ordinary Least Squares Regression Models Predicting Verbal Achievement

	Null Model			Model 1			Model 2		
	Effect	*SE*	*t*	Effect	*SE*	*t*	Effect	*SE*	*t*
Effects of Student Variables									
Verbal achievement score	***27.55	0.04	772.95	***32.74	0.09	370.52	***32.13	0.10	314.16
Number of siblings				***-0.49	0.01	-36.60	***-0.46	0.01	-34.69
Two-parent household				***1.61	0.08	20.27	***1.50	0.08	19.20
Black				***-10.11	0.08	-125.53	***-5.94	0.12	-47.76
Hispanic				***-6.28	0.16	-40.48	***-5.23	0.15	-34.17
American Indian				***-5.85	0.22	-26.55	***-4.45	0.22	-20.06
Asian American				***-1.72	0.28	-6.03	***-1.15	0.28	-4.09
Other				***-9.00	0.15	-59.85	***-7.05	0.15	-45.96
Parents' education				***2.70	0.04	75.84	***2.31	0.04	63.43
Reading material in home				***0.42	0.01	32.97	***0.37	0.01	29.42
Family resources				***0.89	0.04	21.98	***0.37	0.04	8.78
Urbanism of background				***0.84	0.03	26.40	***0.56	0.03	17.07
Effects of School Variables									
Percent Blacks							***-10.21	0.12	-83.28
School mean family resources							***1.51	0.07	20.98
School mean parental education							***3.41	0.05	66.78
South									
Metro									
Science laboratory facilities									
Extracurricular activities									
Comprehensiveness of the curriculum									
Guidance counselors									
Accelerated curriculum									
School enrollment									
Tracking									
Volumes per student									
Movement between tracks									
Expenditure (average teacher salary)									

Average years of experience									
Preference for middle class									
Localism									
Percent White teachers									
Teacher education level									
Family education level									
Verbal score									
Prop families own encyclopedia									
Transfers									
Proportion planning to attend college									
Average hours homework									
Attendance									
R^2				37.13%	39.17%				

(Table 5 continues)

(Table 5 continued)

	Model 3			Model 4			Model 5		
	Effect	*SE*	*t*	Effect	*SE*	*t*	Effect	*SE*	*t*
Effects of Student Variables									
Verbal achievement score	***46.57	6.06	7.68	***32.23	7.21	4.47	*15.69	7.42	2.11
Number of siblings	***-0.46	0.03	-17.61	***-0.47	0.03	-17.77	***-0.46	0.03	-17.63
Two-parent household	***1.64	0.16	10.24	***1.64	0.16	10.23	***1.58	0.16	9.91
Black	***-5.32	0.25	-21.16	***-5.34	0.25	-21.22	***-5.36	0.25	-21.33
Hispanic	***-4.80	0.34	-14.15	***-4.67	0.34	-13.72	***-4.67	0.34	-13.74
American Indian	***-3.74	0.50	-7.46	***-4.08	0.51	-8.07	***-4.52	0.51	-8.87
Asian American	***-4.77	1.15	-4.13	***-4.76	1.15	-4.12	***-4.82	1.15	-4.19
Other	***-6.43	0.33	-19.34	***-6.46	0.33	-19.47	***-6.48	0.33	-19.54
Parents' education	***2.17	0.08	28.57	***2.17	0.08	28.62	***2.17	0.08	28.63
Reading material in home	***0.32	0.03	12.54	***0.32	0.03	12.52	***0.31	0.03	12.18

Family resources	***0.47	0.08	5.68	***0.46	0.08	5.62	***.048	0.08	5.81
Urbanism of background	***0.73	0.08	9.47	***0.73	0.08	9.42	***0.79	0.08	10.03
Effects of School Variables									
Percent Blacks	***-8.52	0.32	-26.78	***-12.18	0.61	-20.01	***-12.32	0.63	-19.58
Family resources	***1.85	0.17	11.15	***1.99	0.18	11.04	***1.82	0.23	7.84
School mean parental education	***3.51	0.11	31.98	***3.28	0.13	25.61	***2.43	0.16	15.12
South	***-1.40	0.24	-5.86	***-1.19	0.26	-4.58	***-1.61	0.27	-6.04
Metro	*-0.49	0.20	-2.43	-0.08	0.22	-0.37	-0.18	0.22	-0.80
Science laboratory facilities	**0.37	0.11	3.21	***0.45	0.12	3.85	*0.25	0.12	2.16
Extracurricular activities	-0.03	0.03	-1.00	*-0.10	0.04	-2.50	-0.07	0.04	-1.82
Comprehensiveness of the curriculum	**-0.19	0.06	-3.47	***-0.24	0.06	-4.06	***-0.20	0.06	-3.30
Guidance counselors	*0.13	0.06	2.24	0.11	0.06	1.88	0.10	0.06	1.68
Accelerated curriculum	0.04	0.07	0.54	0.15	0.08	1.90	0.06	0.08	0.71
School enrollment	-0.25	0.19	-1.32	0.04	0.20	0.21	0.002	0.20	0.01
Tracking	***1.01	0.20	4.97	**0.68	0.21	3.21	***0.62	0.21	2.92
Volumes per student	*0.05	0.02	2.45	**0.06	0.02	2.63	0.04	0.02	1.95
Movement between tracks	***-8.91	1.23	-7.24	***-7.44	1.36	-5.47	***-6.23	1.40	-4.43
Expenditure (average teacher salary)	*-1.49	0.70	-2.14	0.48	0.83	0.57	**2.30	0.86	2.67
Average years of experience				0.03	0.16	0.19	0.02	0.18	0.11
Preference for middle class				0.24	0.16	1.52	0.10	0.16	0.60
Localism				***-0.25	0.08	-3.26	***-0.46	0.08	-5.81
Percent White teachers				***-4.53	0.65	-6.97	***-2.54	0.69	-3.70
Teacher education level				***-0.97	0.26	-3.73	***-1.31	0.27	-4.84
Family education level				-0.24	0.29	-0.81	-0.10	0.30	-0.33
Verbal score				**0.80	0.24	3.33	***0.76	0.24	3.09
Proportion families own encyclopedia							-0.09	0.99	-0.09
Transfers							***-0.96	0.31	-3.10
Proportion planning to attend college							***3.60	0.38	9.52
Average hours homework							0.40	0.34	1.16
Attendance							0.51	0.31	1.63
R^2		41.08%			41.24%			41.50%	

* $p < .05$. ** $p < .01$. *** $p < .001$.

these results, schools from metropolitan areas and from the South performed more poorly than nonmetropolitan schools from other parts of the country. Better resources in terms of access to more guidance counselors, science lab facilities, and library volumes were associated with higher verbal achievement test scores. Schools that had tracking policies had better verbal achievement outcomes, but both a more comprehensive array of curricular options and greater movement between tracks were associated with poorer outcomes. Finally, after controlling for all other curriculum and resource measures, increased expenditures, as measured by the school-level average teacher salary, exhibited a negative relationship with achievement. The standard errors for all these coefficients, though, were underestimated by the OLS model, the hypothesis tests were prone to Type I errors, and these reports of statistically significant outcomes were, thus, specious.

Model 4: Adding teacher characteristics predictors. The introduction of the teacher characteristics in the multilevel Model 4 in Table 4 explained little additional between-school variance in the school mean achievement outcome. After controlling for the facilities and curriculum measures and the teacher characteristics, including average teacher salary, teachers' verbal scores, years of experience, teachers' education levels, and the teachers' family education levels, there was a statistically significant and negative relationship between the school-level percent of White teachers and achievement. In addition, the inclusion of the teacher variables did explain away some of the compositional effects. After controlling between-school differences in teacher characteristics, the coefficient for school mean parental education dropped to less than half its previous magnitude in Model 3 and was no longer a statistically significant predictor of school mean achievement. The introduction of the teacher characteristics did not have the same effect on the school percent Blacks and mean family resources measures, though. In this model, the magnitude of the compositional effect for the school-level percent African American enrollment was 1 3/4 times larger than the individual-level effect of being Black, and the coefficient for school mean family resources was more than 2 1/2 times larger than the student-level family resources coefficient.

The Black slope and the family resources slope were the two other outcomes of Model 4. For both outcomes, we added as a predictor the one teacher characteristic that we hypothesized was associated with within-school social class and Black-White inequalities: the school-level measure of teachers' preference to teach middle-class students over disadvantaged students. Within schools that exhibited stronger teacher preferences to work with middle-class students, the achievement gaps separating Black and White students and students from more and less economically advantaged family backgrounds were amplified. For the statistically significant coefficient of -2.21 for the preference for middle-class measure predicting the Black slope, for example, an increase of 1 standard deviation in teacher bias toward more advantaged students was associated with a widening of the within-school Black-White achievement gap of 0.17 standard deviations.

The OLS model did not include these multivariate outcomes that accounted for both within-school and between-school variability in verbal achievement. A series of interaction terms could provide estimates of cross-level interaction effects, but the error terms would all be negatively biased, and the model would be quite cumbersome given the number of interactions in the slope models. Rather, the teacher characteristics entered as predictors in Model 4 of Table 5 explained only a small fraction of additional variability in the verbal achievement outcome beyond that which was explained by the prior OLS model. Schools whose teachers performed better on a verbal test also helped their students achieve higher verbal achievement test scores. After controlling for teachers' verbal scores and average salary, and other characteristics of the teachers and their schools, schools that had teaching staffs composed of a greater percentage of White teachers, with higher education levels, and who were hired largely from the community in which the school was situated achieved poorer outcomes.

An interesting finding was that, after controlling for the teacher characteristics, the OLS coefficient for the school-level percent Black predictor increased to -12.18. In this model, the compositional effect of the school-level racial/ethic context was 1 1/4 times the magnitude of the student-level effect of being African American. The compositional effect of the school mean family resources was 3 1/3 times that of the individual level effect of the group-mean centered family resources predictor. Finally, similar to the previous OLS models, there was no compositional effect for school mean parental education.

Model 5: Adding student body characteristics predictors. In the final HLM model in Table 4, the five student body characteristics were modeled as school-level predictors of school mean verbal achievement. Only the proportion of students planning to attend college was a statistically significant predictor of the outcome. After controlling for the other student and school characteristics, the model predicted a 3.15-point difference between a school in which all students planned to attend college, and a school in which no students planned to attend college. In this final model, we accounted for nearly 94% of the between-school variance in school mean achievement.

In this final model, even after adjusting the school mean achievement outcome for the objective family background characteristics, school, teacher, and peer effects, the school-level African American and school mean family resources compositional effects were both more than 1 3/4 times the magnitude of the respective student-level effects.[13] These results indicated that the achievement difference between a school with no African American students and a school of 100% African American enrollment was more than 1 3/4 times greater than the achievement difference between an African American student and a White student. Similarly, the achievement difference between a school attended by students of average wealth and a school with a student body composed of students 1 standard deviation below the mean level of wealth was nearly 1 3/4 times greater than the achievement difference between a student of average wealth and a student who was 1 standard deviation less wealthy. Therefore, above and beyond the individual effects of race/ethnicity and poverty, and above and beyond the effects of other school-level resources, there are highly important contextual effects associated with attending more highly segregated schools with higher concentrations of poverty.

The final OLS analysis, Model 5 in Table 5, included the student body characteristics and accounted for only a fraction of 1% of additional variability in the verbal achievement outcome. In this final OLS model, including all student and school predictors, we accounted for 41.50% of the overall variance in verbal test scores. Like the multilevel model, the proportion of students panning to attend college was a statistically significant and positive predictor of achievement. In addition, these results showed a statistically significant negative relationship between the frequency of student transfers and verbal achievement. Similar to the results from the multilevel analysis, the compositional effect of the school-level racial/ethic context was nearly 1 1/3 times the magnitude of the student-level effect of being African American. The compositional effect of the school mean family resources was more than 2 3/4 times that of the student-level effect of family resources. The same technical problems apply to this OLS model as prior models, and thus, these results should be interpreted with caution.

DISCUSSION

Using the original EEO data, this study replicated Coleman's statistical models but also applied a two-level HLM to measure the effects of school-level social composition, resources, teacher characteristics, and peer characteristics on ninth-grade students' verbal achievement. An HLM model was applied because data in education are generally hierarchical in nature. A clear hierarchy consists of students nested within classrooms, and classrooms nested within schools. Analyses that do not take this hierarchy into account produce biased and incorrect results.

HLM explicitly models the nested structure of the data and produces estimates that allow an accurate prediction of outcomes for members of groups as a function of the characteristics of the groups, as well as the characteristics of the members. Most important, the methodology allows researchers to disentangle how schools and students' family backgrounds contribute to learning outcomes. The methodology offers a clearer interpretation of the relative effects of school characteristics, including racial/ethnic composition and family background, on students' academic outcomes. This approach enhances the level of precision in the estimates, thus increasing the quality of inferences made from the data.

In comparison, the traditional OLS regression approach, which Coleman and past analysts of the EEO data employed, applies a single equation and predicts student outcomes at only one level. This causes problems in estimating the variation in achievement outcomes and in turn affects the accuracy of inferences that can be made from the data. When we estimate, for instance, the effects of student and school characteristics in the same equation that predicts student-level achievement outcomes, we are assuming that the school and individual characteristics are from a simple random sample. This is clearly not true because large numbers of individuals were sampled from each of the schools represented in the data set. The school characteristics are all the same for the group of students who are enrolled within the same school. Therefore, the "true" variance in school characteristics is underestimated by OLS. In addition, when clustered or nested data are submitted to a traditional regression analysis, the assumption of independence of units of analysis—a fundamental assumption in statistical analysis—is violated, which leads the researcher to falsely conclude that results are statistically significant and reliable.

In contrast to previous analyses and interpretations of the EEO data, the current analysis focused directly on comparing estimates of the relative achievement effects of a student's family background—race/ethnicity and social class—and the social composition of the school that he or she attended. In addition to these sources of inequality that may be explained by differences among schools, we examined potential within-school sources of inequality within an integrated analytical and theoretical framework. This was accomplished through the application of the two-level multilevel model, which partitions the variance in verbal achievement into its between- and within-school components and models at the appropriate level of aggregation the student and school predictors of the outcomes.

In addition, rather than conducting these analyses on small subsamples of students of particular racial/ethnic backgrounds and from specific regions of the country, the current analysis included the total student and school samples within a comprehensive model. This contemporary approach for analyzing large national data sets, which is common today, has only relatively recently been made possible through the advent of important advances in computing technology. Given the application of these recent methodological and technological developments to the original EEO data, how did it affect our understandings of the relative contributions of families and schools to educational inequality?

First, we find evidence that schools do indeed matter, in that when one examines the outcomes across the national sample of schools, 40% of the variability in verbal achievement is found between schools. Second, even after adjusting for students' family background, a large proportion of the variation among true school means is related to differences that are explained by school characteristics. Third, our multilevel models reveal substantial school-to-school variability in terms of the within-school social distributions of achievement. These within-school inequalities in the achievement outcomes for African American and White students and students from families of higher and lower social class are explained in part by teachers' biases favoring middle-class students and by schools' greater reliance on curriculum differentiation through the use of more comprehensive forms of academic and nonacademic tracking.

Fourth, the results from our OLS regression models reveal stark differences between the traditional specification and conceptualization of school effects, which are akin to the type employed by past analysts of the EEO data, and the contemporary method of modeling school effects through the application of HLMs. Similar to previous reports from Coleman et al. (1966a) and others who reanalyzed the data, the OLS models suggested that school compositional effects, school curriculum and resources, teacher characteristics, and student body characteristics explained very little additional variability in achievement beyond student background—only an additional 4%. In contrast, the HLM models reveal that school composition alone explains nearly one quarter of the variability in the true school means above and beyond student background. However, when one actually compares the magnitudes of the OLS coefficients for an individual's social class status or status as an African American, they are smaller than the school-level coefficients for percent Black and school mean family resources. In this way, the results from the

OLS models and multilevel models are alike. The traditional OLS model, though, would only be capable of estimating the important sources of within-school inequality through a cumbersome series of interaction terms, and it cannot capture the true multilevel conceptualization of the problem. Indeed, the OLS model is not appropriate for disentangling school and individual effects from either a statistical or a conceptual standpoint.

Finally, formal decomposition of the variance attributable to individual background and the social composition of the schools provides very clear and compelling evidence that going to a high-poverty school or a highly segregated African American school has a profound effect on a student's achievement outcomes, above and beyond the effect of his or her individual poverty or minority status. Specifically, both the racial/ethnic and social class composition of a student's school are more than 1 3/4 times more important than a student's individual race/ethnicity or social class for understanding educational outcomes. In dramatic contrast to previous analyses of the Coleman data, these findings reveal that school context effects dwarf the effects of family background.

HOW DOES SCHOOL CONTEXT MATTER?

Although socioeconomic and racial segregation of schools explains a great deal, the relative weakness of the various EEO predictors representing substantive school policies and characteristics suggests that many of the traditional production function measures of school features may be ineffectual or irrelevant for understanding *how* school social context matters. Since the Coleman report, researchers have paid considerable attention to school effectiveness and the practical matter of school improvement (Teddlie & Reynolds, 2000). In many respects, this line of research was a direct reaction to the Coleman report because researchers, including Edmonds (1979), were most intent on demonstrating that schools can and do make a difference for poor and minority children.

Comparatively little research, however, has explored directly the theoretical and practical dimensions of how schooling in high-poverty and racially segregated contexts can restrict students' educational opportunities and outcomes. The question of how the poverty and minority concentration within a school affects a student's achievement outcomes above and beyond the effect of his or her individual poverty and minority status is at the core of the sociology of education. Yet, surprisingly little research has helped develop better statistical estimates and conceptual theories of school compositional effects. To help guide future studies and, ultimately, educational policy decisions, refined and expanded theoretical models are needed. These models must help us understand, among other things, how schools and other institutions, neighborhoods and other forms of collective socialization, and peer effects can contribute to, or mediate, contextual effects.

It is troubling that differences in school resources, teacher characteristics, and student body characteristics help so little in explaining how schools played significant roles in both racial and socioeconomic-based inequality. As Metz (1998) noted, though, on the surface, high schools can look very similar in terms of their architecture and facilities, time schedules, curricular scope and sequence, class sizes, duties for teaching staff, and methods of instruction. From building to building (Meyer & Rowan, 1978) and from decade to decade (Tyack & Tobin, 1994), many aspects of schooling remain remarkably similar and resistant to change. This standardization implied by all schools' adherence to a common script can cover obvious inequities between schools in privileged and disadvantaged contexts (Metz, 1990).

To gain a better picture of inequality, studies of school compositional effects must focus more clearly on that which Metz (1998) termed the "veiled inequalities." Unfortunately, the Coleman report revealed little beyond the common script and surface details of American schools. When we were able to use the EEO data to measure the more subtle forms of inequality that operate, including teachers' biases favoring White and middle-class students and greater within-school curricular differentiation, we were able to demonstrate more clearly how schools exacerbate inequalities. Having further measures of the nature of interactions among teachers and students, of how resources were actually deployed within schools, and of the actual content and quality of classroom instruction would likely help identify other sources of inequality that explain both between- and within-school sources of inequality.

IMPLICATIONS

Coleman et al. (1966a) had originally concluded that the "beneficial effect of a student body with a high proportion of white students comes not from racial composition per se but from the better educational background and higher educational aspirations that are, on the average, found among whites" (p. 306). As Wong and Nicotera (2004) noted, these findings related to the importance of student body characteristics were translated by policy makers and the public, in part for legal reasons and in part for cultural and political reasons, into a discussion of racial integration. The Coleman report was authorized as part of the Civil

Rights Act of 1964 and was conceived of within the context of the legal system's growing reliance on social science to inform legal decisions, most notably *Brown*. During the late 1960s and early 1970s, the Supreme Court endorsed busing to prompt desegregation efforts, which were being implemented at a discouragingly slow rate. Because of White flight and other legal and policy setbacks—including *Milliken v. Bradley*, which ruled against cross-district or metropolitan school desegregation in 1974, and the Emergency School Aid Act, which banned the use of federal funds for busing also in 1974—these efforts were short lived, and school desegregation was significantly curtailed (Salomone, 1986; Wong & Nicotera).

The *Brown v. Board of Education of Topeka* decision stated that even though the physical facilities and other "tangible" factors of White and African American schools may be equal, segregation on the basis of race denies to African American children equal protection under the law. The consistency between the results of our analysis and this central opinion of the court are rather remarkable. Though we find few tangible resources or other factors that explain the effects of school composition, it is clear that racially segregated schools compromised African American students' opportunity to achieve educational outcomes equal to those of their peers at majority-White schools. In addition to the consistency of our results with the recommendations of the Court, they also contradict the conclusions of Coleman and his colleagues, who attributed differences between the outcomes of majority-Black and majority-White schools to the better educational backgrounds and aspirations found among Whites. Our final analytical model, which added statistical controls for a number of student body characteristics—including various measures of educational backgrounds and aspirations—revealed that they explained away none of the effects related to the racial or socioeconomic contexts of schools. In this way, the past misinterpretations by policy makers and researchers that emphasized racial integration over the importance of the student body's educational backgrounds and aspirations are actually very well founded by our contemporary analyses.

Despite these revisionist interpretations of the Coleman report, there are several inherent limitations of the data that cannot be overcome by multilevel models and contemporary computers. We have discussed, for instance, the problems of missing data. In addition, our analyses relied on cross-sectional achievement data and, as a result, were not capable of estimating the growth over time in students' achievement. Longitudinal analyses of the achievement gains made by students across the sampled schools would more accurately represent the potential "value-added" effects of schools. No correlational analyses, though (whether cross-sectional or longitudinal), can support strong causal inferences. Some recent studies, including the Moving to Opportunity Study (Ludwig, Ladd, & Duncan, 2001), which randomly assigned some economically disadvantaged families to receive the assistance they needed to move to more integrated neighborhoods, have offered more convincing evidence of the effects on students' academic outcomes of attending more integrated schools and living in more integrated communities. More such studies of the causal effects of racial and socioeconomic integration are needed, along with complementary descriptive research to document more clearly *how* compositional effects of neighborhoods and schools are manifested.

Since the time of the Coleman report, a substantial research base has grown indicating that children from poverty (Brooks-Gunn & Duncan, 1997) and from African American backgrounds (Jencks & Phillips, 1998) are at considerable risk for poor school performance. Understanding and addressing inequality due to one's social class or racial/ethnic background remains a key societal issue, but the current analysis points to the social context of one's neighborhood and school as a central problem to be confronted by continued theoretical developments, further research, and future social and educational policies. Rather than the all-too-familiar summary of the Coleman report's findings that "schools don't matter," this analysis suggests that both within-school interactions among students and educators, and racial segregation across schools deny African American children equality of educational opportunity.

Acknowledgments

This article was written with the support of a National Academy of Education/Spencer Postdoctoral Fellowship award. Other funding was provided by the Center for Research on the Education of Students Placed at Risk (CRESPAR), a national research and development center supported by a grant (No. R-117-D40005) from the Office of Educational Research and Improvement (OERI), U.S. Department of Education, and by a grant from the U.S. Department of Education, OERI, National Institute on Educational Governance, Finance, Policymaking and Management, to the Consortium for Policy Research in Education (CPRE) and the Wisconsin Center for Education Research, School of Education, University of Wisconsin—Madison (Grant No. OERI-R308A60003). The opinions expressed are those of the authors and do not necessarily reflect the view of the funding agencies, including the National Institute on Educational Governance, Finance, Policymaking and Management, OERI, U.S. Department of Education, the institutional partners of CPRE, or the Wisconsin Center for Education Research.

Notes

1. For a more thorough discussion of the literature on inequality and its relationship to tracking, curricular differentiation, and teacher biases, see, respectively, Oakes (1985), Oakes, Gamoran, and Page (1992), and Ferguson (1998).

2. In the original Coleman report, the regression analyses of school, teacher, and student body effects on achievement controlled only the six objective family background variables. We replicate this approach here because we, like Coleman, are interested in assessing the effects of schools net of family background. The original EEO report and past reanalyses did investigate the relations between achievement and subjective student characteristics, including parents' educational desires and student attitudinal measures, such as control of one's environment. Exploration of these relationships, though, is beyond the scope of this article and not related to our primary objectives.

3. Rather than using a single category, Hispanic, the original Coleman report differentiated between Puerto Ricans and Mexican Americans. Because of relatively small samples of Puerto Rican and Mexican American students, we established the combined Hispanic category.

4. The original Coleman report used the school-level variable percent White rather than percent African American. Because a prominent goal of our research was to measure the magnitude of the compositional effect of attending a highly segregated African American school relative to the individual effect of being African American (see the Procedure section), we chose to use percent African American.

5. Bowles and Levin (1968) were highly critical of the Coleman report's reliance on district per-pupil expenditure averages. As they noted, there is considerable within-district variability in expenditure data, and the limited variation in per-pupil expenditure that is imposed by averaging expenditures over an entire school district results in an understatement of its relationship to student achievement. In addition, as Bowles and Levin argued, data provided by the U.S. Department of Education indicated that about 90% of instructional expenditures are accounted for by teachers' salaries. For these reasons, we believe that using the school-level average teacher salary as an indicator of school expenditures is a reasonable alternative.

6. The South refers to the combination of two regions denoted as South and Southwest. These two regions comprise the following states: Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia, Arizona, New Mexico, Oklahoma, and Texas. The North region comprises the remainder of the United States.

7. The school location variable included seven alternatives that were coded in the survey in the following manner: rural area = 1; small town with a population of 5,000 or less = 2; city of 5,000–50,000 = 3; residential suburb = 4; industrial suburb = 5; residential area of a larger city with a population over 50,000 = 6; and inner part of a larger city with a population over 50,000 = 7. Metropolitan locations included the categories 6 and 7 and nonmetropolitan locations included the remaining categories, 1–5.

8. Some student body measures, including the average number of families with encyclopedias, were primarily exogenous, and others, including the average hours of homework and attendance, were largely endogenous. As such, the latter variables may reflect aspects of school policies and practices as much as qualities of students and their families that existed prior to school. Indeed, some studies have used student homework completion as a measure of a school's academic climate, or "academic press" (Lee & Bryk, 1989; Phillips, 1997).

9. This methodology is akin to contemporary methods for analyzing clustered data from large national education surveys. In the Coleman report, the authors offered an explanation for separating the racial/ethnic groups in the analyses. Specifically, they stated, "When achievement differs as much as it does between these groups, then to analyze the groups together, without controlling for race or ethnicity of the student, would cause any school characteristics highly associated with race or ethnicity to show a spurious relationship to achievement" (Coleman et al., 1966a, p. 311). In our analysis, we address this con-

cern by statistically controlling student race/ethnicity along with the range of other student background characteristics.

10. The classic OLS regression model, though, does not simultaneously model both intercepts and slopes as outcomes. The HLM models, thus, are necessarily different in that they include prediction models for the within-school family resources achievement slope and the Black-White gap, in addition to a model predicting school mean achievement.

11. Compositional effects within the OLS regression models we estimate are specified through the use of a group-mean centering approach and inclusion of both the group-mean centered student characteristic (i.e., family resources, parent education, and African American) and the school-level aggregate of the characteristic. In this case, the compositional effect is the extent to which the magnitude of the school-level relationship differs from the person-level effect.

12. After treating the Black-White achievement gap as a level 2 random effect, a total of 50 schools from the total analytical sample of 226 dropped from our analyses. In these 50 cases, there was no within-school variation in the Black-White achievement gap to measure because the schools comprised an all-Black student body or were attended by no African American students. This result can be seen in Table 3 by comparing the between-school degrees of freedom of 225 for the null model with the degrees of freedom of 175 for Model 1, which included the student race/ethnicity indicators.

13. In addition, results of general linear hypothesis testing, which contrasted the student-level (within-school) coefficients for Black and family resources with the coefficients for the school-level (between-school) coefficients for, respectively, school percent Blacks and school mean family resources, revealed statistically differences. The hypothesis test of the difference between the student-level Black coefficient and the school-level percent Blacks coefficient revealed that the between-school coefficient was larger than the within-school Black coefficient and that this difference was statistically significant, $p < .001$; $\chi^2(2, N = 168) = 195.11$. Similarly, the between-school coefficient for school mean family resources was larger than the within-school family resources coefficient; this difference was also statistically significant, $p < .001$; $\chi^2(2, N = 168) = 31.29$. We obtained similar results for HLM Models 2–4.

References

Armor, D. J. (1972). School and family effects on Black and White achievement: A reexamination of the USOE data. In F. Mosteller & D. P. Moynihan (Eds.), *On equality of educational opportunity* (pp. 168–229). New York: Random House.

Bidwell, C. E., & Kasarda, J. D. (1980). Conceptualizing and measuring the effects of school and schooling. *American Journal of Education, 88,* 401–430.

Bowles, S., & Levin, H. (1968). The determinants of scholastic achievement: An appraisal of some recent evidence. *Journal of Human Resources, 3,* 3–24.

Brooks-Gunn, J., & Duncan, G. J. (1997). The effects of poverty on children. *The Future of Children, 7*(2), 55–71.

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

Burstein, L. (1980). The analysis of multi-level data in educational research and evaluation. In D. C. Berliner (Ed.), *Review of research in education* (pp. 158–233). Washington, DC: American Educational Research Association.

Coleman, J. S. (1987). Families and school. *Educational Researcher, 16*(6), 32–38.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology, 94*(Suppl.), S95–S120.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., et al. (1966a). *Equality of educational opportunity.* Washington, DC: U.S. Government Printing Office.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., et al. (1966b). *Supplemental appendix to the survey on equality of educational opportunity.* Washington, DC: U.S. Government Printing Office.

Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership, 37,* 15–27.

Ferguson, R. F. (1998). Teachers' perceptions and expectations and the Black-White test score gap. In C. Jencks & M. Phillips (Eds.), *The Black-White test score gap* (pp. 273–317). Washington, DC: Brookings.

Gamoran, A., Secada, W. G., & Marrett, C. B. (2000). The organizational context of teaching and learning: Changing theoretical perspectives. In M. T. Hallinan (Ed.), *Handbook of research in the sociology of education* (pp. 37–63). New York: Kluwer Academic/Plenum.

Goldstein, H. I. (1987). *Multilevel models in educational and social research.* London: Oxford University Press.

Guthrie, J. W. (1995). Implications for policy: What might happen in American education if it were known how money actually is spent? In L. O. Picus & J. L. Wattenberger (Eds.), *Where does the money go? Resource allocation in elementary and secondary schools* (pp. 253–268). Thousand Oaks, CA: Corwin Press.

Jencks, C. S., & Mayer, S. E. (1990). The social consequences of growing up in a poor neighborhood. In L. E. Lyn & M. McGeary (Eds.), *Inner-city poverty in the United states* (pp. 111–186). Washington, DC: National Academy of Sciences.

Jencks, C. S., & Phillips, M. (1998). *The Black-White test score gap.* Washington, DC: Brookings.

Jencks, C., Smith, M., Acland, H., Bane, M.J., Cohen, D., Gintis, H., et al. (1972). *Inequality: A reassessment of the effect of family and schooling in America.* New York: Basic Books.

Lee, V. E., & Bryk, A. S. (1989). A multilevel model of the social distribution of high school achievement. *Sociology of Education, 62,* 172–192.

Ludwig, J., Ladd, H., & Duncan, G. (2001). The effects of urban poverty on educational outcomes: Evidence from a randomized experiment. In W. G. Gale & J. R. Pack (Eds.), *Brookings-Wharton papers on urban affairs* (pp. 147–201). Washington, DC: Brookings.

Metz, M. H. (1990). Real school: A universal drama amid disparate experience. In D. Mitchell & M. E. Goertz (Eds.), *Education politics for the new century: The twentieth anniversary yearbook of the Politics of Education Association* (pp. 75–91). Philadelphia: Falmer.

Metz, M. H. (1998, April). *Veiled inequalities: The hidden effects of community social class on high school teachers' perspectives and practices.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Meyer, J. W., & Rowan, B. (1978). The structure of educational organizations. In M. W. Meyer & Associates (Eds.), *Environments and organizations* (pp. 78–109). San Francisco: Jossey-Bass.

Mosteller, F., & Moynihan, D. P. (Eds.). (1972a). *On equality of educational opportunity.* New York: Random House.

Mosteller, F., & Moynihan, D. P. (1972b). A pathbreaking report. In F. Mosteller & D. P. Moynihan (Eds.), *On equality of educational opportunity* (pp. 3–66). New York: Random House.

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

Oakes, J., Gamoran, A., & Page, R.N. (1992). Curriculum differentiation: Opportunities, outcomes, and meanings. In P. W. Jackson (Ed.), *Handbook on research in curriculum* (pp. 570–608). New York: Macmillan.

Phillips, M. (1997). What makes schools effective? A comparison of the relationships of communitarian climate and academic climate to mathematics achievement and attendance during middle school. *American Educational Research Journal, 34,* 633–662.

Raudenbush, S. W., & Bryk, A. S. (1986). A hierarchical model for studying school effects. *Sociology of Education, 59*, 1–17.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage.

Rogosa, D. R. (1978). Politics, process, and pyramids. *Journal of Educational Statistics, 3,* 79–86.

Salomone, R. C. (1986). *Equal education under the law: Legal rights and federal policy.* New York: St. Martin's Press.

Smith, M. (1972). Equality of educational opportunity. In F. Mosteller & D.P. Moynihan (Eds.), *On equality of educational opportunity* (pp. 230–342). New York: Random House.

Teddlie, C., & Reynolds, D. (2000). *The international handbook of school effectiveness research.* London: Falmer Press.

Tyack, D., & Tobin, W. (1994). The "grammar" of schooling: Why has it been so hard to change? *American Educational Research Journal, 31*, 453–479.

Wilson, A. B. (1959). Residential segregation of social classes and aspirations of high school boys. *American Sociological Review, 24,* 836–845.

Wong, K. K., & Nicotera, A. C. (2004). *Brown v. Board of Education* and the Coleman report: Social science research and the debate on educational equality. *Peabody Journal of Education, 79,* 122–135.

GEOFFREY D. BORMAN is a professor of education and sociology at the University of Wisconsin–Madison. His main substantive research interests revolve around the social distribution of the outcomes of schooling and the ways in which policies and practices can help address and overcome educational inequality. Dr. Borman's primary methodological interests include the synthesis of research evidence, the design of quasi-experimental and experimental studies of educational innovations, and the specification of school-effects models. His recent publications include articles appearing in *Review of Education Research*, *Educational Evaluation and Policy Analysis*, and the *American Educational Research Journal.*

MARITZA DOWLING is a faculty member in the Department of Biostatistics and Medical Informatics at the University of Wisconsin–Madison. She holds a master's degree in measurement and statistics and a PhD in quantitative methods. Most of her training has focused on the convergence of quantitative methodology, psychometrics, and cognition. She is particularly committed to bridging the gap between emerging developments and substantive applications of novel analytic methods using cross-sectional and longitudinal data. Her substantive interests include uses and extensions of quantitative models (e.g., multilevel, latent growth, Markov chain, and item-response modeling) to address existing problems in the study of cognitive developmental and stage-sequential processes. Dr. Dowling's recent publications include articles appearing in *Review of Education Research* and *Educational Evaluation and Policy Analysis.*