DOCUMENT RESUME

ED 361 902 EA 025 297

AUTHOR	Kennedy, Eugene; And Others
TITLE	A Multilevel Analysis of Phase II of the Louisiana School Effectiveness Study.
PUB DATE	93
NOTE	26p.; Paper presented at the Annual Meeting of the International Congress for School Effectiveness and Improvement (4th, Norrkoping, Sweden, January 1993).
PUB TYPE	Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Achievement; Educational Environment; Effective Schools Research; Elementary Education; *Models; Multiple Regression Analysis; *Regression (Statistics); *School Effectiveness
IDENTIFIERS	*Louisiana School Effectiveness Study

ABSTRACT

This paper presents findings of a study that used conventional modeling strategies (student- and school-level) and a new multilevel modeling strategy, Hierarchical Linear Modeling, to investigate school effects on student-achievement outcomes for data collected as part of Phase 2 of the Louisiana School Effectiveness Study. The purpose was to identify correlates of achievement, correlates of distribution of achievement within schools, and the stability of school rankings. Data for a variety of social-psychological, experiential, and demographic variables had been collected for 5,400 third-graders, 250 teachers, and 74 principals from a random sample of 76 elementary schools. This study conducted various analyses for norm-referenced subtests in reading, verbal skills, mathematics, language, and a composite. Findings indicate that correlates of achievement and school rankings were consistent across methods. However, the results for the socioeconomic status-achievement slope were less consistent as the strategy moved from the single to the multilevel approach, suggesting that older methods should be viewed with some caution. Finally, mathematics achievement warrants special attention as distinct from more verbal subject areas. Thirteen tables are included. (LMI)

ED 361 902

A Multilevel Analysis of Phase II of The Louisiana School Effectiveness Study

Eugene Kennedy
Louisiana State University

Charles Teddlie
Louisiana State University

Sam Stringfield
Johns Hopkins University

Please send all correspondence to:

Eugene Kennedy
EDAF
111P Peabody Hall
Louisiana State University
Baton Rouge, LA 70803

Telephone: 504-388-6900
FAX: 504-388-2267

Abstract

This paper reports the results of a study which used conventional (student- and school-level) and a relatively new multilevel modeling strategy (Hierarchical Linear Modeling) to investigate school effects on student achievement outcomes for data collected as part of Phase II of the Louisiana School Effectiveness Study. Data for a variety of social psychological, experiential, and demographic variables were collected for students, teachers, and principals for a random sample of 76 elementary schools throughout the state of Louisiana. The study had four objectives: (1) to estimate the impact of schools on student achievement outcomes; (2) to identify school characteristics such as context, and management practices which correlate with student achievement outcomes; (3) to investigate issues of equity with respect to the distribution of student achievement in schools; and (4) to identify schools which appear unusually effective or ineffective at fostering student achievement. Various analyses were conducted for norm-referenced subtests in reading, verbal skills, mathematics, language, and for a composite. The analysis sought to highlight the influence of subject area and methodological considerations for conclusions regarding each of the four issues investigated in the study. The results found mathematics to be the most distinct of the subject areas considered for each of the four objectives considered. Additionally, the correlates of the average achievement within schools identified by the school-level and multilevel modeling strategies were fairly consistent. When the focus was on the distribution of achievement within-schools as indexed by the social class-achievement slope, there was much less consistency. The differences were contributed to the unreliability of the within-school slopes. Finally, Spearman correlations of the rankings of schools with the three strategies were generally high.

Introduction

For several years, educators in the United States (US) and in a number of other countries have sought to identify organizational, social psychological, and management characteristics of schools which correlate with student achievement outcomes (see, for example, Brookover, et al., 1979; Rutter, et al., 1979; Teddlie, et al., 1984; Mortimore, et al., 1988)). From this literature several indicators of effectiveness have emerged and in some states in the US, various indicators have been incorporated in school improvement programs and in general, have been embraced by the educational community (e.g., Segars and Gottesman, 1989). However, the vast majority of this literature is based on research conducted in the late 70s' and early 80s', with reports of large scale comprehensive efforts appearing to have leveled off in recent years. This becomes somewhat problematic given that the regression based residual approaches used in much of this early work has come under increasing criticism. In several critical reviews, investigators have argued that the student-level and school-aggregate regression approaches used in much of this literature will typically yield less satisfactory parameter estimates than recently proposed multilevel analyses strategies. Additionally, critics have argued that multilevel modeling provides a much more direct means of addressing various issues in effective schools research and will often yield different results from those of the student and school regression based approaches (Aitkin & Longford, 1986; Raudenbush & Bryk, 1988).

Although a number of significant applications of multilevel modeling have appeared in recent literature, especially with regards to the High School and Beyond Survey data (Lee & Bryk, 1989), there appears to be lacking, a systematic, substantively comprehensive effort to investigate the implications of model selection for key issues in school effects research. Previous studies which have focused on model selection have typically incorporated empirical data for illustrative purposes, with little substantive import. The study reported in this paper is one of the first substantively comprehensive efforts along these lines. In particular, this study will investigate the impact of model selection (student, school, and multilevel) and subject area criterion (reading, language, mathematics, verbal skills, and a composite) on conclusions regarding four key issues in effective schools research: (1) estimation of the impact of schools on student achievement outcomes, (2) identification of school characteristics such as context, and management practices which correlate with student achievement outcomes; (3) issues of equity with respect to the distribution of student achievement; and (4) identification and comparison of schools which are unusual at fostering student achievement.

It should be noted that this study is not intended to address the question of whether multilevel modeling or the older single level regression strategies are appropriate for various issues in effective schools research. Several comprehensive reviews have made a strong case for the multilevel strategy, although its' practical utility for some school effects issues has been questioned (May, 1990). Rather, the purpose of the current investigation was to provide empirical insight into to how studies which have employed these older techniques should be viewed in light of recent technical advances. It has been observed, for example, that because multilevel modeling allows for a distinction between sampling variability and parameter variability, investigators in the past could have been misled about the true strength of a relationship because parameter estimates were contaminated by large amounts of sampling error (Raudenbush & Bryk, 1986). Conversely, several examples of dramatic reversals in conclusions as analysis moved from a single to a multilevel strategy, have been reported in the literature (Aitkin & Longford, 1986). The current study, by reexamining a large-scale ongoing project will provide some insight along these lines.

The paper is organized into four parts. Part one describes the Louisiana school effectiveness project which is the source of the current data. This section also reviews issues of model selection and subject area criterion as it relates to the key issues addressed in the study. Part two describes the data base actually used in these analyses. It is noted that the results reported in this study may differ from previously published reports of this study due to the deletion of certain students. Parts three and four present the results and conclusions,

The Louisiana School Effectiveness Study

The Louisiana School Effectiveness Study (LSES) is one of the more systematic and comprehensive attempts in America to explore the dynamics of school effects on student achievement outcomes. The project was divided into several phases. Phase I involved conceptualization and instrument development; Phase II involved analysis of data for a sample of 76 schools; and as part of the remaining phases of the project a select group of schools have been followed for nearly ten years at this writing (Teddlie et al., 1984). For the current study, data from the LSES Phase II are utilized. Using interviews, questionnaires, and other data gathering techniques, data were collected on a vast array of variables for more than 5400 students, 250 teachers, and for each of the principals in the 76 schools studied. Data elements include socioeconomic background and other demographic information of students, teachers, and principals; information on faculty experience, training, and interactions with students and the administration; student, teacher, and principal expectations, perceptions of classroom structure and activities, self concept, locus of control and other social, experiential, and psychological variables.

As with many of the landmark studies of effective schools, previously published reports of the LSES Phase II, because multilevel modeling was not developed to the point of practical application at the time, are based on a school-level regression analysis with a norm-referenced composite score as the criterion (see Teddlie et al., 1984). These analyses were designed to replicate and expand upon those presented by Brookover et al. (1979) in their landmark multivariate study of school effectiveness in Michigan. Toward this end, the large number of data elements (nearly a hundred items on each of the teacher and principal questionnaires) included in the study were factor analyzed separately for school socioeconomic indicators, items from the student, teacher, and principal questionnaires. Dominant factors which were significantly related to achievement were retained for additional analyses. These analyses sought to identify correlates of achievement and to identify schools which, given SES factor scores, yielded unusually high levels of achievement.

Models of School Correlates of Achievement

More than 20 years have passed since Dyer, Linn, & Patton (1969) recommended that conclusions regarding school effects on student learning be based on the "discrepancy" between actual school mean achievement and the achievement expected given previous achievement and the social background of the students. As implied, the proposed strategy was simply to aggregate all student-level data to the school-level and using multiple regression, school characteristics linked to achievement could be identified. The identification of unusual schools could be based on the standardized discrepancy between actual achievement and the achievement predicted on the basis of school context and the composition of the student body.

Criticisms of the school aggregate residual model include: (1) parameter estimation is generally inefficient, (2) problems of multicollinearity are common, (3) school effectiveness indices based on this model are generally unstable, and (4) the focus on means may mask important differences among subgroups of students with regard to school effects (Aitkin & Longford, 1986; Seltzer, 1987). With respect to the instability issue, Seltzer (1987) has presented a particularly revealing link between instability and the reliability of the residual from the regression model. Basically, because the regression strategy formulates the expected performance for schools from the conditional distributions of the predictors, each of which will have restricted parameter variability, the resultant residuals tend to be heavily influenced by sampling error. For these and other reasons, both Aitkin and Longford (1986), Raudenbush and Bryk (1988), as well as Seltzer (1987) conclude that this model is likely to yield untrustworthy results for either identifying correlates of achievement (note the ecological fallacy and aggregation bias) or identifying unusual schools.

As an alternative to the school aggregate model, a number of investigators interested in identifying unusual schools employ the student-level residual model. Accordingly, students from all schools are pooled together and, with indifference to school, the criterion of interest is regressed on student background and other student variables considered relevant. The residuals from this total sample regression are then averaged by school and taken as an index of school effectiveness. However, it can be shown that this strategy will yield biased estimates of school effects if there is heterogeneity of slopes among schools. Despite this obvious shortcoming, this strategy is currently used in some statewide school award systems and has been considered preferable to the multilevel strategy discussed below (May, 1990).

From a design standpoint, the question of school effects on student outcomes is inherently a multilevel data problem. It can be said of the school aggregate and student-level models mentioned above, that they represent polar strategies for dealing with difficulties associated with the nesting of students within schools. In lieu of these strategies, a variety of procedures have been proposed over the years under such names as contextual models, and slopes as outcomes (Burstein, 1980). Only recently, however, have advances in statistical modeling and estimation allowed researches to formulate and evaluate models which realistically reflect the hierarchical nature of educational data. Under the general names of multilevel models, variance component models, and hierarchical models, these procedures allow researchers to model complex within-school processes as consequences of between-school characteristics such as context, climate and management practices (see Mason, Wong, & Entwisle, 1983; Raudenbush & Bryk, 1986; Goldstein, 1986; Aitkin & Longford, 1986).

Following the Raudenbush & Bryk (1986) presentation, within each of the schools in the study, the achievement of student i in school j (Y_{ij}) is considered a function of student background characteristics, and a host of experiential and social psychological variables. Thus within each school we have:

$$Y_{ij} = B_{jo} + B_{j1}X_{ij1} + - - - + B_{jk}X_{ijk} + R_{ij} \qquad (1)$$

where B_{jk} is the regression coefficient associated with variable X_k in school j, and R_{ij} is random error assumed to be normally distributed with mean 0 and finite variance. It is assumed that B_{jo} as well as some of the B_{jk} terms will vary across schools and that this variation will be related to school policies, practices and/or environment. For each of the parameters in Equation 1 we will therefore pose a between-school model. That is,

$$B_{jk} = Z_{ok} + Z_{1k}W_{1j} + - - - + Z_{pk}W_{pj} + U_{jk} \qquad (2)$$

where the W terms represent school characteristics, the Z terms represent coefficients relating school characteristics to within-school slopes, and U_{jk} is random error assumed to be normally distributed with mean 0 and finite variance.

Questions regarding the correlates of student achievement in this formulation can not only be posed for mean levels of achievement within each school, but additionally, between school models of the within-school slopes, particularly those associated with student background characteristics, are informative about school characteristics linked to the distribution of achievement in schools.

An additional advantage of the multilevel formulation over the school-level model is the distinction between the error terms associated with Equations 1 and 2. While the R_{ij} term in Equation 1 is interpreted as the usual random error component associated regression analysis, the U_{jk} term in Equation 2 represents the difference between the magnitude of the actual within-school parameter B_{jk} in school j and the magnitude expected given the between-school characteristics **W**. This distinction allows for a separation of the variance of the B_{jk} terms into sampling variance and parameter variance. It follows that between-school characteristics can be related to both, total observed variation of within-school parameter estimates and to estimated parameter variation of the within-school parameter estimates. This distinction is not possible in the school-level model.

In the context of multilevel modeling, the problem of the identification of unusually effective or ineffective schools can be address in terms of the difference between the achievement expected given student characteristics in school j and the achievement expected given aggregate characteristics of school j. This difference can be expressed in terms of the Ujk terms in Equation 2 (see Tate, 1988). In particular, the uniqueness of school j is given by,

$$U^* = Uj0 + - - - - + UjkXjk \quad (3)$$

Because of the presence of student variables in Equation 3, the effects of school j can be investigated for students with different characteristics. That is, with this model it is not necessary to assume that the effect of school j is constant for all students.

Typically, the residuals in Equation 3 are based on school context and composition variables. However, Raudenbush and Bryk (1988) have argued that such models are biased. Analytically, they show that models of school effects which ignore those factors which produce effectiveness are generally biased. As opposed to U* they propose the follow definition of a school effect:

$$U^{**} = Uj0 + ZW + - - - + (ZWjk + Ujk)^*Xijk \quad (4)$$

where W represents those school characteristics (climate, management practices, etc.) which lead to differential effectiveness.

Focusing on the empirical Bayes (EB) estimates associated with multilevel modeling, May (1990) has questioned the utility of these procedures for identifying unusual schools. His concerns are based on the shrinkage phenomenon associated with the EB estimates wherein within-unit parameters are shrunken toward the conditional mean in proportion to their sampling error. Because sampling error tends to be a function of the number of observations within units, as Seltzer (1987) notes, if unusual schools tend also to be small, it is unlikely that EB procedures would ever recognize them as unusual-generally their effects would be shrunken toward the conditional mean.

Methodology

Population and Sample

The data for the current study are from Phase II of the Louisiana School Effectiveness Study. A stratified random sample of 76 elementary schools in Louisiana constituted the sample. Data were collected during the 1982-83 school year. Questionnaires designed to elicit opinions about school climate, information on social psychological variables, and respondent background were administered to principals, teachers, and students. Other data were collected from records at district offices and the Louisiana State Department of Education. As note above, the total sample consists of approximately 5400 third grade students, 250 teacher, and 74 principals.

Instrumentation.

The student questionnaire (46 items), teacher questionnaire (99 items) and principal questionnaire (78 items) were factor analyzed as a means of data reduction. Retaining only those factors with eigenvalue greater than 1.0, principal axis extraction with varimax rotaion yielded ten factors from the student questionnaire, 21 factors from the teacher questionnaire, and 17 factors from the principal questionnaire. For the teacher and principal solutions, only the ten factors with the largest eigenvalues were retained. The factors are as follows:

Student Questionnaire

S_EDEXPRE	Students' Present Educational Expectations and Comparisons with Others
S_EDEXFUT	Students' Future Educational Expectations
S_POSCLMT	Positive School Climate
S_TCHPUSH	Teachers' Work and Push
S_NEGCLMT	Negative School Climate
S_CAREGRD	Students and Teachers Care About Grades
S_WKCNTRL	Work Independently and Positive Locus of Control
S_WKHARD	Students Work Hard
S_NEGSLF	Negative Self Image
S_SCHLRN	Learning That Occurs in School

Teacher Questionnaire

T_COLLEXP	College Expectations for Students
T_STDABIL	Student Academic Abilities
T_TRYHARD	Students Try Hard
T_PRNHELP	Principals' Help
T_PREPEXP	Years of Experience and Preparation
T_HSEXP	High School Expectations for Students

T_MTHDATT	Teaching Methods and Attitudes
T_CLSWHL	Class Works as a Whole
T_TCHSLF	Teachers' Self Concept
T_STDSLF	Priority for Enhancing Students' Self Concept

Principal Questionnaire

P_EDEXP	Future Academic Expectations for Students
P_SCHSUC	School Success and Students' Academic Ability
P_PARCAR	Parents' Concern About Grades and Education
P_HRSWK	Hours Spent Working
P_PRNTCH	Principal Working with Teacher
P_CNTRL	Principals' Attitudes and Locus of Control
P_YRSEXP	Years of Experience
P_ATTEND	Presence of Teachers and Principal
P_PRNSELF	Principals' Self Concept
P_PARSUP	Parental Support

Factor analysis also yielded two school level socioeconomic/contextual factors linked to achievement. The first factor (SESHOME) was composed of parents education level and occupation. The second factor (SESRACE) was largely composed of items which indicated the proportion of faculty and students identified as white.

The achievement indicators used in the current study are norm-referenced instruments designed by Scholastic Testing Service (STS) for use in Phase II of the LSES project. The instruments are shortened versions of the Educational Development Series (EDS), lower primary level achievement tests developed by STS. The instruments were administered to the students in the sample from January to March of 1983. The measurement consists of a composite (BSK) and subtests in the subject areas of Reading (RD), Verbal Skills (VB), Mathematics (MT), and English (ENG). Reviews of the instrument by the Louisiana Department of Education found it to be consistent with the curriculum and instructional objectives of the state.

Additional information on Phase II of the LSES can be found in Teddlie et al. (1984).

Analysis Strategy

The regression analyses described in this study were conducted with procedures found in the SAS statistical package. The multilevel modeling analyses was accomplished with a computer program developed by Bryk, Raudenbush, Seltzer, and Congdon (1986). The program generates a residual file which can be used to assess the tenability of the normality assumption for the between model error terms. The so called Q-Q plots were examined for the various analyses presented below, and did not show any drastic departures from normality.

Missing Data

Data for the principal factors in the between-school model were missing for two schools. Estimates were obtained by regressing these measures against all other predictors in the between-school model. This means of handling missing data in the between model will likely yield better estimates then simply plugging in sample means. Missing data for the within-school model was handled by pairwise deletion.

Results

Schools and Student Achievement

This part of the study was concerned with the proportion of variation in student achievement which could be linked to the schools that students attend. If between school differences in achievement were minimal, then subsequent investigation of the data with respect to schools would have questionable utility. However, significant between school variation would be cause to investigate characteristics of the student population, as well as school management practices, climate and other attributes. In analysis of variance (ANOVA) parlance, the relevance of schools can be assessed by partitioning the total variation in student achievement into its' within and between-school components. For this purpose the VARCOMP routine in the Statistical Analysis System (SAS) computing package was used to estimate the variance components for a oneway Model II ANOVA with schools as a random factor. These results and the associated intraclass correlation coefficients (RHO) are presented in Table 1 for the four subtests and composite used in this study.

Table 1
Estimates of Variance Components

Source	BSK	ENG	MT	RD	VB
School	69.91	10.11	3.50	12.86	10.18
Error	455.02	72.62	50.42	98.62	70.49
RHO	0.13	0.12	0.06	0.11	0.12

These results are consistent with others in that there appears to be considerably more variability in student achievement within-schools than between-schools (Mandeville & Anderson, 1986). With the exception of mathematics (MT), school to school differences in achievement account for approximately 12% of the observed variation in achievement. While this indicates that the bulk of student achievement variation for these data occurs within-schools, it is important to note that these measures provide little information about the potential of schools to impact student achievement. Small intraclass correlations would be observed if, assuming no student selection effect, all schools were equally effective or ineffective at fostering student achievement. For this reason, albeit the measures are small, it is important to proceed with the investigation toward the end of identifying school characteristics which may have an impact on achievement.

Because the between-school differences in student achievement reported in Table 1 are certainly affected by random sampling error, it is of interest to estimate that proportion of school to school variation which reflect true differences. The multilevel analysis strategy employed in this study can address this issue in that it is possible estimate the proportion of the observed variation in school mean achievement that is parameter variability. Additionally, it is possible to estimate the proportion of the observed variation among schools in the link between student achievement and socioeconomic background that is parameter variation. In this, the distribution of achievement within schools is investigated with respect to school characteristics, which, given the large variances reported in Table 1, would certainly warrant attention.

For these analyses we again consider the within-school model in Equation 1 wherein the achievement of student i in school j (Y_{ij}) is considered a function of various student-level characteristics. For the analysis considered in this report, we limit consideration to a within-school model which posses achievement as a function of student socioeconomic background (SES), defined as a composite of parents education and occupation. Accordingly,

$$Y_{ij} = B_{j0} + B_{j1}X_{ij} \qquad (5)$$

where B_{j0} is the intercept in school j and B_{j1} is the slope of the regression of achievement on SES in school j. To further interpretation of these results, within each school, student SES is rescaled by deviating it from the average SES for that school. In this way the slope and intercept parameters can be interpreted as follows:

B_{j0} = (the average achievement in school j)

B_{j1} = (the differentiating effect of SES background).

To estimate the proportion of the variation among schools in the within-school slope and intercept estimates that is true parameter variation, we again consider the between-school model depicted in Equation 2. In this instance, however, the within-school parameters are considered a function of the average parameter for the population of schools only. Thus,

$$B_j = M_B + U_j \qquad (6)$$

where M_B represents the average B parameter for the population of schools considered, and U_j is error assumed to be normally distributed with mean 0 and finite variance. This variance, Var(U), is in effect the parameter variance associated with the parameter in question. The ratio of the estimate of parameter variance to observed variance for the slope and intercept terms in the within-school model provide an indication of the proportion of the observed variation that is parameter variation (frequently called the reliability of the estimates). Table 2 presents the results for this model, normally referred to as the "unconditional model", for the five measures of achievement considered in this study. In addition to the reliability of the slope and intercept estimates, this analysis also tests the hypotheses that the average slope and intercept parameters are zero, and that the parameter variability associated with these terms is not significantly different from zero. This latter hypothesis is critical in that logically, only parameter variability is potentially explainable.

Table 2
HLM Results for the Unconditional Model

	BSK	ENG	MT	RD	VB
Mean School Achievement	97.62	30.60	39.04	27.97	32.61
Mean SES-Achievement Slope	5.45	2.20	1.23	2.08	2.18
Variation in Mean School Achievement					
Parameter Variation	72.89	11.23	4.26	12.64	9.83
Observed Variation	82.35	12.70	5.37	14.64	11.29
Reliability	0.88	0.88	0.79	0.86	0.87
Variation in School SES-Achievement Slopes					
Parameter Variation	12.81	1.38	1.45	2.08	1.09
Observed Variation	31.90	4.35	3.64	6.11	4.03
Reliability	0.40	0.31	0.39	0.34	0.27

Note. All parameter estimates are significant at the 0.001 level.

The results presented in Table 2 clearly indicate that the vast majority of the variation in mean achievement among schools is parameter variability. As expected, the slope estimates tend to have more sampling error but nevertheless, a respectable proportion of the observed variability is parameter variability. Additionally, the average effect estimate for the slopes indicates that for the population of schools considered, student socioeconomic background (SES), as expected, tends to have a positive effect on student achievement.

It is worth noting that this analysis provides an estimate of the true correlation between the slope and intercept terms for the various subject areas. Thus the estimated true correlation between the SES-achievement slope and mean achievement is 0.13 for BSK, 0.11 for ENG, -0.45 for MT, 0.41 for RD, and -0.05 for VB. The large negative coefficient for mathematics suggest that in schools where the link of math achievement with home background is significantly positive, the average achievement level of the school in this subject area is relatively low.

The coefficients in the table are all in the expected directions. Schools with students from high SES backgrounds; where a large percentage of the students and faculty are white; where the climate promotes academic achievement and concern for grades; and where administrators and faculty hold high expectations for student achievement and work toward improving academic performance, are schools which yield higher levels of average achievement. These results are similar to those of other effective schools studies and reiterate those of previously published reports of this project (see Levine & Lozotte, 1990; Teddlie et al. 1984).

Although the school-level regression approach has been utilized chiefly to address issues of the correlates of achievement and the identification of unusual schools, it is possible to investigate correlates of the distribution of achievement. To do so, the ordinary least squares (OLS) estimate of the SES-achievement slope within each school was obtained. As with school mean achievement, stepwise regressions were run with the 32 school factors constituting the pool of potential predictors. These results are presented in Table 4 for each of the five achievement indicators studied.

Table 4.
Regression Results for School OLS Slopes

	--------- Subject Area ---------				
	BSK	ENG	MT	RD	VB
R^2	(0.22)	(0.07)	(0.15)	(0.24)	(0.15)
Predictor					
SESRACE	****	****	****	0.21	****
S_WKCNTRL	-0.30	****	****	-0.28	****
S_NEGSLF	****	****	-0.21	****	****
T_STDSLF	0.21	0.26	****	****	0.30
P_YRSEXP	0.31	****	0.30	****	0.25
P_PARSUP	****	****	****	0.28	****

Note. Table entries are standardized Beta Weights.

The large negative coefficient for S_WKCNTRL reported in Table 4 indicates that schools where students typically work independently and possess a locus of control which promotes responsibility for their own actions, are also those where the link of socioeconomic background is less likely to determine achievement. Additionally, two other indicators related to student self perceptions (S_NEGSLF and T_STDSLF) yield significant coefficients. Without including these measures in the within-school models, it is difficult to assess their mediating effect. However, the results do suggest the importance of attitudinal factors at mediating background influences on achievement.

As preparation for the multilevel analysis of these data it was again desirable to reduce the number of potential predictors. Toward this end, Pearson correlations of the school factors with the EB residuals for the intercept and slope of the unconditional

School Correlates of Achievement

This part of the study was concerned with the identification of characteristics of schools related to student achievement outcomes. In previous analyses of these data the school-level aggregate model was employed for this purpose using mean student achievement as the outcome of interest. In the current study we contrast these results with those obtained from an application of the HLM methodology. Because many of the early studies of school effectiveness employed the school-level aggregate model, this contrast is particularly interesting.

Because of the large number of potential predictors, Pearson correlations of each achievement indicator with the 32 factors from the student, teacher, and principal questionnaires were obtained. The pool was then reduced to those factors which yielded significant simple correlations. These factors were then entered into a stepwise regression wherein the criterion for retention in the model was a probability value smaller than 0.05. The results for each of the achievement measures is presented in Table 3.

Table 3
Regression Results For School Aggregate Model

	Subject Area				
	BSK	ENG	MT	RD	VB
R2 (SES)	(0.56)	(0.57)	(0.27)	(0.58)	(0.71)
R2 (Total)	(0.75)	(0.72)	(0.62)	(0.69)	(0.77)
Predictor					
SESHOME	0.41	0.44	****	0.51	0.50
SESRACE	0.20	0.21	****	0.26	0.26
S_EDEXFUT	0.20	0.13	0.23	0.23	****
S_NEGCLMT	0.17	0.18	0.26	****	****
S_CAREGRD	0.32	0.30	0.47	0.25	0.23
T_HSEXP	-0.12	****	-0.19	-0.11	****
T_STDABIL	****	****	****	****	-0.16
P_HRSWK	****	****	****	****	0.13
P_ATTEND	****	****	-0.16	****	***

Note. Table entries are standardized Beta Weights.

Table 3 presents multiple regression results for the school-level aggregate model for each of the five achievement indicators studied. From these results the various models appear effective at predicting mean school achievement. With the exception of mathematics, this appears to be true whether all predictors or just the SES predictors are included in the model. For mathematics, the relative predictive power of the average school SES composites have relatively little predictive power. These results are consistent with others in which home background has been found to have generally small relevance for mathematics achievement compared to achievement in verbal areas (Mandeville & Anderson, 1986).

models (see Table 2) were obtained for each of the achievement measures studied. This is similar to the exploratory analysis in the HLM computer program wherein residuals from fitted models are regressed on between-unit predictors not in the model. For current purposes, stepwise regressions were run for the slope and intercept residuals for each subtest against those factors with significant Pearson correlations for the criterion considered. Nine factors emerged from these analyses. The Pearson correlations of these factors with empirical Bayes residuals from the unconditional model are presented in Table 5 for the intercept and in Table 6 for the slope.

Table 5
Pearson Correlations of Empirical Bayes Residuals for Intercept with Factors in Between-School Model

Factors	BSK	ENG	MT	RD	VB
SESHOME	0.66	0.68	0.37	0.71	0.73
SESRACE	0.33	0.37	0.12*	0.36	0.45
S_EDEXFUT	0.28	0.23	0.25	0.29	0.22
S_CAREGRD	0.55	0.55	0.50	0.46	0.47
S_WKCNTRL	-0.19*	-0.23	0.01*	-0.26	-0.29
S_NEGSLF	-0.13*	-0.19*	-0.02*	-0.10*	-0.09*
T_STDSLF	-0.01*	0.00*	-0.15*	0.04*	0.02*
P_HRSWK	-0.15*	0.15*	0.04*	0.20*	0.26
P_YRSEXP	0.01*	0.03*	-0.12*	0.06*	0.03*

* Not statistically significant at the 0.05 level.

Table 6
Pearson Correlations of Empirical Bayes Residuals for SES-Achievement Slope with Factors in Between-School Model

Factors	BSK	ENG	MT	RD	VB
SESHOME	0.43	0.40	0.10*	0.56	0.73
SESRACE	0.25	0.17*	0.09*	0.36	0.45
S_EDEXFUT	0.20*	0.22	0.01*	0.23	0.02*
S_CAREGRD	0.18*	0.17*	-0.10*	0.28	0.01*
S_WKCNTRL	-0.32	-0.26	-0.19*	-0.37	-0.13*
S_NEGSLF	-0.21*	-0.11*	-0.22	-0.16*	-0.13*
T_STDSLF	0.24	0.26	0.21*	0.19*	0.26
P_HRSWK	0.04*	0.09*	-0.03*	0.11*	-0.01*
P_YRSEXP	0.21*	0.14*	0.28	0.15*	0.19*

* Not statistically significant at the 0.05 level.

While there are some differences, many of the indicators in these tables are the same as those found relevant in the school-level aggregate results presented above. While this anticipates some similarities between the school aggregate results and HLM results, the large number of nonsignificant coefficients, especially for the slope residuals, suggests some important departures. The fitted model from the hierarchical linear modeling (HLM) program are presented in Table 7.

Table 7
HLM Results for School Factor Scores

	BSK	ENG	MT	RD	VB
School Mean Achievement					
SESHOME	3.66	1.54	0.54	1.86	1.57
SESRACE	2.36	0.90	****	1.02	1.13
S_CAREGRD	3.11	1.23	0.97	0.88	0.79
S_EDEXFUT	2.46	0.73	****	0.99	0.74
P_HRSWK	****	****	****	****	0.64
SES-Achievement Slope					
SESHOME	1.88	0.57	0.46	0.86	0.47
STDRACE	****	****	****	-0.66	****
Percent Observed Variation Explained					
School Mean Ach.	58.27%	60.70%	27.18%	60.65%	67.93%
SES-Ach. Slope	11.37%	8.96%	10.71%	19.96%	8.43%
Percent Parameter Variation Explained					
School Mean Ach.	65.83%	68.65%	34.27%	70.25%	78.02%
SES-Ach. Slope	28.33%	28.26%	26.89%	53.65%	31.19%

First, as expected, the model is much more relevant to variation in mean achievement than the SES-achievement slope. The amount of parameter variation in means explained is generally twice that of the slope. Additionally, while the fitted model for means in these analyses is similar to that for the school aggregate regression analysis with respect to the predictors present, the results for slopes differ drastically. In the current results, only the SES factors appear to have any relevance to slope variability. Finally, the uniqueness of mathematics is found in both analyses.

An alternative to reducing the number of between-school factors described above would have been to factor analyzed the original factors. These analyses were conducted with the 12 initial factors which yielded significant Pearson correlations with mean achievement for the composite BSK. These factors were; S_EDEXFUT, S_TCHPUSH, S_NEGCLMT, S_CAREGRD, T_COLLEXP, T_STDABIL, P_EDEXP, P_SCHSUC, P_PRNTCH, and P_PARSUP. Four factors emerged from the analysis. These were; Student SES (SESFAC2), Current Academic Climate (CLIMATE), Student Expectations/Parental Support

(EXPECT), and School Academic Caring and Success (SCHLSUC) (see Teddlie et al., 1984).

Table 8
HLM Results for Second-Order Factor Analysis

	BSK	ENG	MT	RD	VB
School Mean Achievement					
SESFAC2	4.01	1.51	0.59	1.90	1.75
CLIMATE	2.76	1.13	****	1.16	1.45
EXPECT	4.40	1.83	1.13	1.44	1.29
SCHLSUC	2.73	0.95	0.76	1.21	0.89
SES-Achievement Slope					
CLIMATE	****	****	****	0.67	****
SCHLSUC	1.31	****	****	0.67	****
Percent Observed Variation Explained					
School Mean Ach.	53.51%	56.37%	26.75%	55.69%	60.47%
SES-Ach. Slope	3.63%	****	****	13.47%	****
Percent Parameter Variation Explained					
School Mean Ach.	60.43%	63.76%	33.79%	64.30%	69.17%
SES-Ach. Slope	8.88%	****	****	38.28%	****

The HLM fitted model for the second-order factor analysis are presented in Table 8. In this case there is more consistency across subject areas for mean variation with respect to the relevant factors. However, the SES-achievement slope reveals considerably more variation in relevant predictors across achievement measures than was the case for the first-order factor scores. Only the composite and the reading subtest yield significant predictors for the slope.

Identification and Comparison of Unusual Schools

As noted above, investigators have employed a variety of strategies for the identification of schools which appear unusually effective or unusually ineffective at fostering student achievement. In this portion of the paper we compare the results obtained from the student and school based residual strategies with that of the multilevel modeling strategy. As noted above multilevel modeling allows for the issues of school effectiveness to be addressed with respect to students of various characteristics. In the current context, school effectiveness indices are obtained for students of low, middle and high socioeconomic backgrounds. To accomplish this the SES variable in the within-school model was rescaled so that each students' score represents a deviation from the average SES for the entire sample. In this way, the intercept of the within-school model is no longer simply the average achievement level in the school, but rather, it becomes the average achievement level in the school adjusted for the SES level of students within the school (see Tate, 1988; Raudenbush & Bryk, 1988).

As with previous results, comparisons of the student residual (STDRESID), school residual (SCHRESID), and multilevel modeling (MM_HI, MM_MID, MM_LOW) strategies for identifying unusual schools will be made for each of the five achievement indicators considered in this study. These analyses should be informative of the degree to which the identification of unusual schools will vary depending upon the method and subject area considered.

Table 9 presents Spearman correlations of the rankings for the composite (BSK) obtained with each of the five methods considered. Table 10 presents similar results for the subtest in mathematics (MT). Comparisons of results among the three multilevel modeling rankings are informative about the degree to which schools serve students across the SES spectrum. For both the composite and math achievement, the rankings for middle and low SES students are more similar to one another than either is to the rankings for high SES students. The distinction in mathematics appears to be greater than that for the composite. Considering rankings across methods, the table indicates that the student residual, school residual and multilevel results for the middle and low SES student are in fair agreement with one another. The rankings associated with the multilevel results for the high SES student appear to be the most distinct of all five results considered. This level of similarity in rankings among the various procedures is encouraging for those investigators who have used the older methodologies to compare schools. However, because most studies are concerned with schools on the extremes of distributions, it is of interest to investigate consistency in rankings for schools at the top and bottom of the distributions. These results are presented in Table 11.

Table 9
Spearman Correlations of Composite (BSK) Rankings Across Method

	SCHRESID	STDRESID	MM_HI	MM_MID	MM_LOW
SCHRESID	1.00				
STDRESID	0.77	1.00			
MM_HI	0.60	0.77	1.00		
MM_MID	0.72	0.82	0.86	1.00	
MM_LOW	0.73	0.78	0.72	0.96	1.00

Table 10
Spearman Correlations of Math (MT) Rankings Across Method

	SCHRESID	STDRESID	MM_HI	MM_MID	MM_LOW
SCHRESID	1.00				
STDRESID	0.93	1.00			
MM_HI	0.57	0.64	1.00		
MM_MID	0.85	0.91	0.66	1.00	
MM_LOW	0.81	0.86	0.46	0.95	1.00

Table 11 presents the percent of schools consistently identified among the top 15 or bottom 15 for each of the methods considered for the composite achievement measure. The lowest values in the table occur for the percent of schools consistently identified in the top 15 percent. In both instances, the disagreement involves the multilevel results for students with relatively high socioeconomic backgrounds. The results for the middle and low SES student rankings appear much more consistent with other results in the table.

Table 11
Percent of Schools Consistently Identified Across Methods
Results for Composite (BSK)

Top 15% \ Bottom 15%

	SCHRESID	STDRESID	MM_HI	MM_MID	MM_LOW
SCHRESID		75	66	83	75
STDRESID	54		50	66	58
MM_HI	45	63		58	50
MM_MID	63	81	63		91
MM_LOW	63	72	45	81	

Table 12 also presents results on consistency of extreme rankings, but in this instance, the rankings occur across subject areas. Because of space considerations, only the results for the school aggregate model and multilevel middle SES rankings are considered. In general, the results indicate much more consistency for the multilevel rankings. In several instances, the school aggregate rankings agree across subject area for only a third of the schools considered. As might be expected, given other results, the rankings for mathematics seem generally in disharmony with the other measures.

Table 12
Percent of Schools Consistently Identified Across Subject Areas
Results for School Aggregate and HLM (Mean SES) Models

Results for School Aggregate Model

Top 15% \ Bottom 15%

	BSK	ENG	MT	RD	VB
BSK		58	41	83	33
ENG	81		41	41	25
MT	54	54		33	16
RD	63	54	36		33
VB	54	45	36	54	

(Table 12 continued)

Results for HLM (Mean SES) Model

Top 15% \ Bottom 15%

	BSK	ENG	MT	RD	VB
BSK		83	66	83	75
ENG	81		75	66	66
MT	63	81		58	50
RD	90	72	54		66
VB	90	81	63	81	

The final issue addressed in this study concerns correlates of the school residual indices. These correlations are presented in Table 13 for the composite achievement measure. The relevant factors are similar to those identified in the earlier analyses seeking to find correlates of achievement. The significance patterns differ as one moves from one method to another. All appear to bear some relationship with mean achievement, especially the student residual approach.

Table 13
Pearson Correlations of School Effect Residuals (BSK) with Factor Scores

Factor	SCHRESID	STDRESID	MM_HI	MM_MID	MM_LOW
SESHOME	NS	.51	.46	.38	.33
SESRACE	NS	.32	NS	NS	NS
T_STDABIL	NS	-.42	-.22	-.24	-.23
S_EDEXFUT	.30	.27	.60	.39	.28
S_NEGCLMT	.33	.56	.39	.35	.31
S_CAREGRD	.42	.58	.69	.57	.48
P_EDEXP	NS	-.23	-.27	-.25	-.23
P_SCHSUC	-.19	-.45	-.31	-.29	-.26
P_PARCAR	NS	-.25	NS	NS	NS
P_PRNTCH	NS	.29	NS	NS	NS
P_ATTEND	NS	-.27	NS	NS	NS
P_PARSUP	NS	.23	.37	.27	NS
MEAN	.66	.95	.75	.73	.67

Conclusions

This study was concerned with differences in substantive conclusions regarding issues in effective schools research which could be linked to the statistical model chosen or the subject area studied. Because practical applications of multilevel analysis procedures have only recently become available, many of the large-scale comprehensive effective schools projects of the early 80s' and late 70s' used student and school aggregate regression strategies which have, of late, come under significant criticisms. Because these efforts have had a profound impact on current activities in effective schools research, it is important to determine if significant differences in conclusions result once multilevel analysis strategies are employed. Toward this end, data from Phase II of the Louisiana School Effectiveness Study was reanalyzed using a student, school, and multilevel analysis strategy. The study sought to identify correlates of achievement, correlates of the distribution of achievement within schools, and the stability of school rankings.

The results regarding correlates of achievement and school rankings were very consistent across methods. Conversely, the results for the SES-achievement slope were less consistent as the strategy moved from the single to the multilevel approach. The difference was attributed to the larger sampling error associated with the slope. This result suggest that while previous reports of correlates of achievement are likely to be replicated with current methods, results regarding correlates of the distribution of achievement obtained with these older methods may wisely be viewed with some caution. Finally, the data support the contention that mathematics achievement warrants special attention as distinct from the more verbal subject areas.

References

Aitkin, M. & Longford, N. (1986). Statistical modeling issues in school effectiveness studies. Journal of the Royal Statistical Society, Series A, 149 (1), 1-43.

Brookover, S., Schweitzer, J., Schneider, J.M., Beady, C.H., Flood, P.K., & Wisenbaker, J. (1979). Elementary school social climate and school achievement. American Educational Research Journal, 15 (2), 301-318.

Bryk, A.S., Raudenbush, S.W., Seltzer, M., & Congdon, R.T. (1986). An introduction to HLM: Computer program and users' guide. University of Chicago, Department of Education.

Burstein, L. (1980). The analysis of multi-level data in educational research and evaluation. Review of Research in Education, 8, 158-233.

Dyer, H.S., Linn, R.L., & Patton, M.J. (1969). A comparison of four methods of obtaining discrepancy measures based on observed and predicted school system means on achievement tests. American Educational Research Journal, 6 (4), 591-605.

Goldstein, H. (1986). Multilevel mixed linear model analysis using iterative generalized least squares. Biometrika, 73, 43-56.

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

Mason, W.M., Wong, G.Y., & Entwisle, B. (1983). Contextual analysis through the multilevel linear model. (Pp. 72-103) in Sociological Methodology 1983-1984, edited by S. Leinhardt. San Francisco: Jossey-Bass.

May, J. (1990). Real world considerations in the development of an effective school incentive program, Paper presented at the annual meeting of the American Educational Research Association, Boston, MA.

Mortimore, Peter, et. al. (1988). School matters. Berkley: University of California Press.

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. (1988). Qualitative models for estimating teacher and school effectiveness. Paper presented at the Invitation Conference of the Center for Student Testing, Evaluation, and Standards, Princeton, New Jersey.

Rutter, M., Maughan, B., Mortimore, P. & Ouston, J. with Smith, A. (1979). Fifteen thousand hours: Secondary schools and their effects on children. London: Open Books.

Segars, J.K. & Gottesman, B.L. (1989). The reliability and validity of the effective schools needs assessment surveys. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Seltzer, M.H. (1987). An empirical Bayes approach to the identification of effective schools. Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.

Tate, R.L. (1988). School comparisons with hierarchical linear models: An illustration. (Technical Report). Tallahassee, FL: Florida State University.

Teddlie, C., Falkowski, C., Stringfield, S., Desseelle, S., & Garvue, R. (1984). Louisiana School Effectiveness Study: Phase Two 1982-84 (Research Report). Baton Rouge, LA: Louisiana State Department of Education.