DOCUMENT RESUME

ED 410 243 TM 026 980

AUTHOR	Von Secker, Clare E.; Lissitz, Robert W.
TITLE	Estimating School Value-Added Effectiveness: Consequences of Respecification of Hierarchical Linear Models.
PUB DATE	Mar 97
NOTE	34p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 24-28, 1997).
PUB TYPE	Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Achievement; Accountability; *Educational Practices; *Effective Schools Research; *Estimation (Mathematics); Grade 10; High School Students; High Schools; Institutional Characteristics; Performance Factors; *School Effectiveness; Socioeconomic Status; Student Characteristics; Suburban Schools; Urban Schools
IDENTIFIERS	*Hierarchical Linear Modeling; *Value Added Model

ABSTRACT

The last two decades have seen an increase in the demand for information about school effectiveness. For this paper, school effectiveness is defined as the value-added effectiveness. Value-added indices measure achievement over and above what would have been expected from attendance at a comparable school. They assess school effectiveness after controlling for contributory variables such as socioeconomic status or pre-existing student characteristics. This study examined variation among schools using a hierarchical linear model (HLM) to test whether factors that have a theoretical basis are capable of capturing the unique contribution of specific school practices and policies to student achievement. Data were drawn from the 1992 High School Effectiveness Study, part of the second wave of the National Education Longitudinal Study of 1988. The sample size was 7,642 students, representing 790,810 tenth graders in 247 urban and suburban schools in the largest metropolitan school districts. At least one teacher completed a questionnaire for 5,228 of these students, and at least one teacher answered a questionnaire in 207 of the schools. Findings suggest that widely accepted beliefs about the strength of the associations of school characteristics with student achievement need re-evaluation. Four results seemed robust to methodological considerations: (1) student-level characteristics matter, and are invariant to school policy; (2) the willingness of a school to control problems is strongly associated with performance; (3) schools should not expect stronger leadership, clearer goals, and stricter graduation requirements to improve student performance; and (4) efforts to improve performance by modifying instruction and holding teachers more accountable for student outcomes are probably wasted. An appendix discusses student and school characteristics. (Contains 12 tables and 33 references.) (SLD)

Estimating School Value-Added Effectiveness: Consequences of Respecification of Hierarchical Linear Models

Clare E. Von Secker
Robert W. Lissitz

This paper is prepared for the Annual Meeting of the American Educational Research Association in Chicago, IL

ESTIMATING SCHOOL VALUE-ADDED EFFECTIVENESS: CONSEQUENCES OF RESPECIFICATON OF HIERARCHICAL LINEAR MODELS

Clare E. Von Secker and Robert W. Lissitz
University of Maryland

INTRODUCTION

The last two decades have seen a marked increase in the demand for information about school effectiveness (Orsak, et al., 1996; Raudenbush & Willms, 1995; Webster & Mendro, 1995; Kennedy, 1992). Concerns about standards, accountability, and student achievement have sparked debate about valid ways of measuring and reporting the extent to which schools influence what students know and are able to do. Increasingly, states and local school districts are focusing on differences in performance of comparable students at different schools, and ranking schools based on perceived effectiveness. Accompanying this focus is controversy over how to identify effective schools.

For purposes of this study, school effectiveness is defined as value-added effectiveness, an interpretation suited to analysis through hierarchical linear modeling. Value-added indices are those that measure achievement over and above what would have been expected from attendance at a comparable school. Value added indices seek to "level the playing field" by assessing school effectiveness after controlling for contributory variables that transcend school sovereignty, such as socioeconomic status, and pre-existing student characteristics such as prior achievement and motivation; they examine the impact of school policies and programs on student achievement.

Literature on school effectiveness (Webster, et al., 1994; Dow and Oakley, 1992; Purkey and Smith, 1983) consistently identifies characteristics of effective schools. They include clear, well-defined goals; strong leadership; high expectations for students; and positive, orderly school climate. Defining these constructs operationally, and collecting data to determine the extent to which these factors are present in a school, is time-consuming and expensive. While there is some reason to believe that student achievement is greater, on average, in schools that possess these qualities than it is in schools without them, consistent findings across time and subject areas have not been verified (Mandeville, 1988; Mandeville and Anderson, 1987).

One limitation associated with measuring the association of school effectiveness indicators and student achievement can be the assessments used to evaluate student outcomes. Critics argue that outcome measures tend to be limited to performance on standardized tests of basic skills that measure only part of a multidimensional construct (Mandeville and Anderson, 1987; Mackenzie, 1983; Rowan, et al., 1983). They argue that achievement measures that fail to include assessment of higher order thinking skills may underestimate school effects.

Also, the ability of schools to influence student achievement may be greater in some areas than in others. The total amount of free variation, that is, the extent to which schools may potentially have impact on performance, seems to fluctuate across subject matter. For example, Mandeville and Anderson (1987) reported that South Carolina elementary schools showed more free variation in mathematics than in reading. Findings by Sammons, et al. (1993) corroborate a claim that school effects account for more variance in mathematics achievement. In general, correlation among cognitive tests areas is modest ($r = .61$), leaving plenty of room for differences in statistical outcomes.

Another limitation is that overall school effectiveness is typically represented by performance of a small sample of students in one or two subject areas. Even large-scale studies may provide unrepresentative samples if the amount of missing data is substantial. Studies suggest that correlations of student achievement across subjects and academic years range from modest to low (Crone, et al., 1995; Mandeville & Anderson, 1987; Mandeville, 1988; Mackenzie, 1983; Rowan, et al., 1983). Measures of school effectiveness thus present an incomplete picture of student achievement within a school. Cross-validation studies using sub-samples of students across subject areas may help us estimate the extent to which this limitation biases interpretation of associations of school effectiveness indicators with student achievement.

One well-studied indicator of school effectiveness is school climate. However, interpretation of the results of school climate studies is limited by inconsistency among studies (Lee, et al., 1996; Anderson, 1982). One overarching problem is that while school climate is associated with school effectiveness, definitions of school climate vary greatly and include a range of factors, including school organization, student attitudes and behaviors, school orderliness and safety, and academic learning environment. The extent to which differences in the focus of the definition affect the association of this indicator with student achievement is unknown.

RESEARCH QUESTIONS

Our study was organized around three research questions:

1. To what extent are school effectiveness indices sensitive to changes in measures of achievement?

2. How are inferences about the strength and direction of associations of school value-added indices and school achievement affected by differences in sample size and composition?

3. How does specification of the construct "school climate" affect interpretation of its association with schools effectiveness?

Our study examined variation among schools using a hierarchical linear model, HLM, to test whether factors that have a theoretical basis are capable of capturing the unique contribution of specific school practices and policies to student achievement. A hierarchical model allows us to control for differences due to student characteristics and due to school characteristics that are resistant to changes in policy. In our models, variation associated with prior student achievement, student expectations, and student socioeconomic status was controlled before school level characteristics were considered. In addition, two school level factors, a measure of school socioeconomic status (percent of students receiving free and reduced meals), and school sector (public or private) were controlled when other value-added school effectiveness indices were evaluated. Five value-added indices measured the contribution of variables that can be manipulated at the school level. They were:

- The extent to which school goals and policies are clear
- The strength and effectiveness of school leadership
- The degree to which disruptive climates interfere with instruction
- The academic expectations for students
- The accountability of teachers to facilitate student progress

Current versions of HLM estimate the results of regressions on one dependent variable only. That means that simultaneous assessment of the associations of value-added indices with reading, mathematics, science, and history is not possible. Therefore, we used a series of models that differed only in the subject in which achievement was measured to compare relationships between school effects and achievement across four subject areas. Our aim was to determine whether achievement in reading, mathematics, science, and history were equally sensitive to differences in school policies and characteristics. If the amount of free variation across subject areas were significant, changes in the specification of the hierarchical models could affect the nature of the inferences that might be drawn about these school effects.

Our second goal was to determine the extent to which our models were sensitive to two different methods of handling missing data: replacing missing values with means versus listwise deletion of cases. HLM places a number of restrictions on missing data. The researcher must choose between deleting cases where data is missing, or substituting some value, such as a mean, for missing data. We used two models, each of which measured achievement in four subject areas. The first included all original cases, but missing values on every variable were replaced with the mean for that variable. The second model deleted all cases for which a value was missing on one or more variables.

Finally, we investigated the consequences of respecifying the construct "school climate." One potential problem with evaluation of school effectiveness is that different studies use different measures to define the same construct. We examined the sensitivity of our models to changes in the definition of this construct. Our five definitions captured different aspects of school climate:

- academic learning environment

- school attendance
- student behavior and attitudes
- orderliness and discipline
- violence and crime

The Sample

Data for this study were drawn from the 1992 High School Effectiveness Study (HSES), conducted as part of the second wave of the National Education Longitudinal Study of 1988 (NELS:1988), sponsored by the National Center for Educational Statistics (NCES), U. S. Department of Education. The High School Effectiveness Study was designed specifically to facilitate nested analyses of the type discussed in this paper. Detailed information about design and analysis of the NELS:88 database is available in the High School Effectiveness Study: Data File User's Manual (Scott, et al., 1995).

The sample size is 7,642 students, representing 790,810 tenth grade students enrolled in 247 urban and suburban schools in the 30 largest metropolitan school districts. School level information came from two sources: teachers and administrators. Teachers completed a four-part questionnaire that asked them to describe the type of class the student was in, the kinds of instructional and curricular choices they make, their background and training, and an evaluation of school climate. At least one teacher completed a questionnaire for 5,228 of the 7,642 students. Teacher responses were averaged across all students in a school. At least one student had a teacher who answered a questionnaire in 207 of the 247 schools. Administrator questionnaires were available for each of the 207 schools that had teacher data.

In order to use a data file in HLM, we had to delete cases and schools that were missing data on any of the variables we planned to use in our two-level model. When we accounted for missing data from students, teachers, and school administrators, our working database was reduced to 3449 cases in 136 schools. We constructed a second database to restore the original number of cases and schools by replacing all missing values with the mean for their respective items. These two models are identified in this paper as the Listwise Deletions model and the MEANS model, respectively. Evaluation of these two methods of handling missing data was one of the goals of this study.

Dependent Variable Specification

Tenth Grade Academic Achievement. HSES measures achievement in four subject areas: mathematics, science, reading comprehension, and history (history /citizenship/geography). Mathematics test items include word problems, graphs, equations, quantitative comparisons, and geometric figures. Science questions were drawn from the fields of biology, earth science, physics, and chemistry. The reading comprehension subtest contains five short reading passages or pairs of passages with three to five questions about the content of each. History test items address issues and events in American political and economic history, the working of the federal government and rights and obligations of citizens, and patterns of settlement and food production. The tests were developed by the Educational Testing Service (ETS) with the express purpose of measuring higher order thinking skills as well as understanding of fundamental concepts and basic skills. Reports on the psychometric properties of the cognitive tests can be obtained from NCES.

INDEPENDENT VARIABLE SPECIFICATION

Student Characteristics. The HSES Data Users Manual recommends including specific student background *critical items* in hierarchical models of school effectiveness. Our preliminary analysis suggested that our student-level model could sustain three composite factors. We included (1) a composite of student expectations about high school graduation and pursuit of advanced degrees (EXPECT); (2) a composite indicator of student prior academic performance and achievement (PERFORM); and (3) a socioeconomic composite variable (SES) prepared by NCES. Level-1 predictors were expected to partial out achievement variance that reasonably could be associated with differences among student bodies at different schools. Unlike school effectiveness measures, these variables are attributes that are beyond the control of the school and represent the "what students bring to school" characteristics. A list of items used for the composites is presented in the appendix.

School Context Variables. The HSES Data Users Manual also identifies a number of school context variables as *critical items* for assessing school effectiveness, including school size, number of teachers, length of school day, and percent of students receiving special services. Two factors, school sector (PUBSCH) and percent of the student body eligible for free and reduced means (PCNTFARM), were selected as context variables for these models because preliminary analysis of a range of critical context items suggested that these were significant predictors of school level variance. School context characteristics are included to address differences in the make-up of the student body that cannot be manipulated by school policy, yet affect the achievement of students.

School Policy Variables. School policy factors isolate the effects of school policies on student achievement. Linear composites of equally weighted variables were developed using principal components analysis of the school level data to measure constructs associated with school effectiveness. LEADRSHP is a teacher report of the extent to which the principal and other administrators demonstrate strong leadership within the school. GOALSCLR is a teacher report of the extent to which the goals and mission of the school are clear. Measures of high expectations were divided into two composites: teachers and schools. TQUALITY reflects how much teachers report holding themselves accountable for facilitating student progress, particularly with groups of students who have trouble mastering subject matter in the teacher's particular area. REQGRAD measures how much schools hold students accountable for demonstrating mastery of knowledge and skills before conferring diplomas, namely, whether or not students are required to pass minimum competency tests in English, mathematics, science and history, and the number of credits required for graduation. CLIMACAD measures the extent to which teachers report that maintaining a positive, orderly, academic climate is a problem within the school.

In addition to the climate composite CLIMACAD, four alternative variables were used to capture different facets of school climate. The composite CLIMVICR measures the extent to which violence and crime are a problem within the school. A single-item variable, CLIMRLEN assesses overall school discipline, or the extent to which school rules are consistently enforced. The composite CLIMATND reflects students' attitudes about attending class by summing the extent to which class cutting, tardiness, and absenteeism are problems within a school. The composite CLIMBEHV identifies the extent to which students behave in an academic and serious manner at school. These alternatives were used to demonstrate how differences in operationalization of a construct can affect inferences about the significance of school effectiveness indicators. A list of items used for constructing each alternative climate variable is included in the appendix.

The unstandardized means and standard deviations of all variables are included in Table 1. SES is a standardized composite that has a mean of 0 and a standard deviation of 1. When missing values for this variable are replaced with means, the mean increased to .20 and the standard deviation was reduced to .81. All of the variables used in the hierarchical models were standardized to have a mean of 0 and a standard deviation of 1. In the hierarchical models, coefficients for these variables can be interpreted as the change in student achievement expected for one standard deviation change in the variable.

TABLE 1: UNSTANDARDIZED MEANS AND STANDARD DEVIATIONS FOR WITHIN-SCHOOL AND BETWEEN SCHOOL INDEPENDENT VARIABLES

	MEANS MODEL (7642 CASES IN 247 SCHOOLS)		LISTWISE DELETIONS MODEL (3449 CASES IN 136 SCHOOLS)	
VARIABLE	MEAN	SD	MEAN	SD
WITHIN SCHOOL				
SES (STANDARDIZED)	.20	.81	.34	.79
PERFORM	3.31	.92	3.38	0.99
EXPECT	10.45	2.22	10.73	2.06
BETWEEN SCHOOL				
CLIMACAD	6.07	.94	5.99	1.06
CLIMATND	2.65	.66	2.59	.71
CLIMBEHV	3.47	.77	3.41	.85
CLIMRLEN	3.31	.84	3.38	.86
CLIMVICR	1.77	.44	1.72	.48
GOALSCLR	4.23	.64	4.25	.71
LEADRSHP	4.14	.58	4.20	.60
PCNTFARM	19.55	24.18	17.12	24.40
PUBSCH	.70	.45	1.33	.47
REQGRAD	11.53	1.79	11.36	2.05
TQUALITY	4.21	.27	4.20	.32

Description of Hierarchical Linear Modeling

The methods cited most frequently in literature over the last ten years are those that analyze residuals from a regression of current achievement on combinations of past achievement, student level variables, and school difference variables. Contributions of schools to student achievement are highlighted after controlling for "hard to change" variables such as differences in racial composition, SES, and prior achievement. Using this procedure facilitates examination of differences in school organization and policies that might account for differences in achievement among schools after differences in student bodies are adjusted to be the same. HLM also allows researchers to estimate the proportion of variance in residuals that might be explained by school context and process variables compared to that which is due to sampling error. Application of this model is providing deeper understanding of the processes of schooling and determinants of school achievement. A highly readable first introduction to HLM is presented in Arnold (1992).

Hierarchical linear modeling (HLM) is a method of explaining differences among individuals who are members of nested groups. One practical application of HLM is investigating educational outcomes such as student achievement, where levels of the hierarchy include students nested within classrooms, classrooms nested within schools, schools nested within school districts, and so forth. HLM provides a means of predicting how variables at a higher level can affect the models for student performance at a lower level. The two-level model used in this study examines differences in achievement among tenth grade students nested within high schools.

At the first level of analysis, a series of regression equations (one per school) predicts student achievement as a function of student characteristics within each school. The *within-unit* regression equations vary as a function of average achievement scores and the relative strength of the effect of student level variables. These varying intercept and slope coefficients are used as dependent variables in the second level equation, where *between-unit* regression equations use school characteristics as independent variables to explain coefficient differences among schools. Thus, HLM can be thought of as a series of regressions on regressions.

Model Specification

We used a two-level hierarchical linear model to evaluate the value-added impact of school policy variables on student achievement in four subject areas when pre-existing student characteristics were controlled. Although HLM tolerates a limited number of missing values in the within-school variables, missing values in the between-school equations are not allowed. Since the number of missing values was substantial, we replaced all missing values for variables with the mean value of that variable across all schools. We used this MEANS model to estimate the effects of student and school variables on achievement in reading, mathematics, science, and history.

The initial MEANS model serves as a baseline measure for comparison with models in the second part of this study. Here, the *within-school* equation for each school consists of an intercept (B_{0j}) that represents the average achievement of students within that school, and regression coefficients (B_{1j}, B_{2j}, and B_{3j}) that estimate the effect of student level variables on achievement, plus sampling error (R_{ij}). The level-1 covariates SES, PERFORM, and EXPECT, (X_{1ij} - X_{3ij}, respectively), were centered around their grand means so that intercepts and regression coefficients were adjusted for differences among schools on these factors. Grand mean centering reduces estimation bias that might result if other significant predictors were not specified in this model (Bryk and Raudenbush, 1992). The predicted achievement score (Y_{ij}) for the i^{th} student in the j^{th} school is:

$$Y_{ij} = B_{0j} + B_{1j}(X_{1ij} - X_{1..}) + B_{2j}(X_{2ij} - X_{2..}) + B_{3j}(X_{3ij} - X_{3..}) + R_{ij}.$$

In the second level of analysis, intercepts and coefficients from the level-1 equation are allowed to vary randomly as the between-school parameters are examined. The number of *between-school* equations is equal to the number of random factors. In this study, the intercept and three coefficients are allowed to vary freely, so the number of level-two equations is four.

One advantage of using HLM rather than multiple regression is that the former model uses within-school variance to estimate parameter variance between schools. Parameters are weighted by the inverse of the standard errors of their within-school estimates so that coefficients from schools with smaller samples and/or greater standard deviations are given less weight. This procedure increases the precision of the variance estimates. Prior to regression on school level variables, each parameter was tested using an intercept-only *unconditional model* that provided weighted averages of each parameter so that significant differences in effects among schools could be detected. Each equation in the unconditional model consists of an intercept (G_{p0}) that represents the average within-school parameter value, and an error term ($U_{p.}$) that represents total random error associated with that parameter. The *between-school* equations for average achievement (B_0) and the average effects of SES (B_1), PERFORM (B_2) and EXPECT (B_3) across *j* schools are:

$$B_0 = G_{00} + U_0$$
$$B_1 = G_{10} + U_1$$
$$B_2 = G_{20} + U_2$$
$$B_3 = G_{30} + U_3$$

The values Wp_1 - Wp_7 in the following equations represent the level-two variables PUBSCH, PCNTFARM, CLIMACAD, LEADRSHP, GOALSCLR, TQUALITY, and REQGRAD. In the initial study, policy variables are expected to explain deviations in average school achievement across schools and why socioeconomic status, past performance, and student expectations affect student achievement more in some schools than in others. All variables were centered around the grand mean ($W_{p.}$). For every school (j), the between-school *conditional* model has the following equations:

$$
\begin{aligned}
B_{0j} = G_{00} &+ G_{01}(W_{01j} - W_{01.}) + G_{02}(W_{02j} - W_{02.}) + G_{03}(W_{03j} - W_{03.}) + G_{04}(W_{04j} - W_{04.}) \\
&+ G_{05}(W_{05j} - W_{05.}) + G_{06}(W_{06j} - W_{06.}) + G_{07}(W_{07j} - W_{07.}) + U_0 \\
B_{1j} = G_{10} &+ G_{11}(W_{11j} - W_{11.}) + G_{12}(W_{12j} - W_{12.}) + G_{13}(W_{13j} - W_{13.}) + G_{14}(W_{14j} - W_{14.}) \\
&+ G_{15}(W_{15j} - W_{15.}) + G_{16}(W_{16j} - W_{16.}) + G_{17}(W_{17j} - W_{17.}) + U_1 \\
B_{2j} = G_{20} &+ G_{21}(W_{21j} - W_{21.}) + G_{22}(W_{22j} - W_{22.}) + G_{23}(W_{23j} - W_{23.}) + G_{24}(W_{24j} - W_{24.}) \\
&+ G_{25}(W_{25j} - W_{25.}) + G_{26}(W_{26j} - W_{26.}) + G_{27}(W_{27j} - W_{27.}) + U_2 \\
B_{3j} = G_{30} &+ G_{31}(W_{31j} - W_{31.}) + G_{32}(W_{32j} - W_{32.}) + G_{33}(W_{33j} - W_{33.}) + G_{34}(W_{34j} - W_{34.}) \\
&+ G_{35}(W_{35j} - W_{35.}) + G_{36}(W_{36j} - W_{36.}) + G_{37}(W_{37j} - W_{37.}) + U_3
\end{aligned}
$$

Schools with higher, positive values for the variables LEADRSHP, GOALSCLR, TQUALITY, and REQGRAD and lower values for the variables PUBSCH, PCNTFARM, and CLIMACAD were expected to have higher average achievement scores (B_{0j}). More desirable scores on these variables were also expected to close the gap in achievement attributed to the student's socioeconomic status, previous performance, and expectations. For example, B_{1j} is average deviation in achievement that can be predicted by knowing the SES of students within a school; the level-2 equation suggests how much variance in achievement due to socioeconomic status is mitigated by school context and policy measures W_{11j}-W_{17j}. Likewise, the equation for B_{2j}, the deviation in achievement that can be predicted by knowing an individual's performance, predicts the extent to which the gap in achievement among students of varying performance is minimized by variables W_{21j}-W_{27j}. Finally, the B_{3j} equation estimates the extent to which school policy variables W_{31j}-W_{37j} can regulate the strength of the association of student expectations about their future academic success with achievement.

Although Arnold (1992) recommends replacing missing values with means, this procedure will reduce the amount of variance around each parameter that can be estimated by school effects. When the amount of missing data is substantial, as was the case in this study, replacing missing values with means may obscure associations of school effects with differences in achievement between schools. Our second group of comparisons provided estimates of the consequences of choosing this alternative. We repeated our initial study of school effects using a model in which cases missing data for any variable were deleted. The *within-* and *between-school* equations for these two models are identical, but the degrees of freedom are not. Our Listwise Deletions model included data for only 136 of the 244 schools in the MEANS model.

A third series of analyses looks at how changing operational definitions of one construct affects the apparent magnitude and direction of seven school effects on student achievement. As with our initial set of comparisons, missing values were replaced with the mean value for the relevant variable across all schools. Since our findings on previous analyses suggested that the school effects we selected did not explain a significant proportion of variance in the slope coefficients, the model was simplified. Variance in average school achievement was estimated as before, but slope coefficients were allowed to vary randomly. The values W_{01} - W_{07} are the same as in the previous between-unit model, *except* that W_{03} is used to represent four different climate variables

CLIMVICR, CLIMRLEN, CLIMBEHV, and CLIMATND. All level-2 variables were adjusted by their grand means ($W_{op.}$). The adjusted between-unit model has the following equations:

$$
\begin{aligned}
B_{0j} = {} & G_{00} + G_{01}(W_{01j} - W_{01.}) + G_{02}(W_{02j} - W_{02.}) + G_{03}(W_{03j} - W_{03.}) + G_{04}(W_{04j} - W_{04.}) \\
& + G_{05}(W_{05j} - W_{05.}) + G_{06}(W_{06j} - W_{06.}) + G_{07}(W_{07j} - W_{07.}) + U_0 \\
B_{1.} = {} & G_{10} + U_1 \\
B_{2.} = {} & G_{20} + U_2 \\
B_{3.} = {} & G_{30} + U_3
\end{aligned}
$$

Expectations about the effects of school level variables for LEADRSHP, GOALSCLR, TQUALITY, REQGRAD, PUBSCH, and PCNTFARM are the same as for the previous between-school model. As with CLIMACAD, high values for variables CLIMATND, CLIMVICR, and CLIMBEHV were expected to correlate with lower academic achievement. Unlike other climate variables, high values for CLIMRLEN were expected to correlate with higher achievement.

Results

Analysis of the Consequences of Using Alternative Measures of Achievement .

The average association of each within-school predictor with student achievement in four subject areas is shown in Table 2. Five of the 246 schools were dropped automatically by HLM because they contained insufficient numbers of students to support the analysis. This unconditional model summarizes the regressions for the remaining 241 schools. Since each parameter in the unconditional model is predicted using only an intercept, the fixed effects coefficients represent the change in achievement in each school associated with one standard deviation change in the variable. The t-tests indicated that the variance around each parameter was significant, although, as expected, the greatest amount of variance was around the parameter B_{0j} in all subject areas.

Overall, these effects are robust measures of achievement in different subject areas. A student's prior performance is the best indicator; on average, an increase in student achievement from 2.24 to 3.03 points is expected for every standard deviation increase in PERFORM. SES appears to be a better predictor of achievement in science and history than in reading and mathematics. Student expectations about graduating high school and pursuing advanced degrees (EXPECT) are more strongly associated with achievement in the "basic" subjects, reading and mathematics, than in science or history.

TABLE 2: Comparison of Average Within-School Predictors of Achievement in Reading, Mathematics, Science, and History using MEANS models									
Unconditional Models		**Reading**		**Math**		**Science**		**History**	
Fixed Effects									
Student Level	School Level	Coeff	p	Coeff	p	Coeff	p	Coeff	p
Intercept B0									
	G00 intercept	49.24	.000	49.80	.000	48.94	.000	49.32	.000
SES B1									
	G10 intercept	1.51	.000	1.41	.000	1.70	.000	1.93	.000
EXPECT B2									
	G20 intercept	1.61	.000	1.62	.000	1.21	.000	1.40	.000
PERFORM B3									
	G30 intercept	2.24	.000	3.03	.000	2.91	.000	2.54	.000
Variance Components (df = 240)									
		Var.	p	Var.	p	Var.	p	Var.	p
	Tau U0	11.65	.000	17.19	.000	15.87	.000	16.22	.000
	Tau U1	2.38	.000	1.90	.000	2.93	.000	3.15	.000
	Tau U2	3.19	.000	2.96	.000	2.48	.000	2.54	.000
	Tau U3	2.94	.000	3.51	.000	3.26	.000	3.18	.000
σ^2 R		25.82		24.83		25.96		28.66	

The between-school results are presented in Table 3. The conditional models used seven school effects to try to explain the variation, or Tau value (U0, U1, U2, and U3), around each parameter. The standardized school context variable for PCNTFARM explained a significant amount of variance in average achievement in all subjects. On average, student achievement scores are predicted to decrease by .70 to .94 points for every one percent increase in the number of students in the school who receive free or reduced meals. Once this school context variable is controlled, however, school sector does not explain any significant variation in average school achievement.

Once school context variables are controlled, only two of the five policy variables emerged as significant predictors of differences in school achievement. The standardized school policy variables for GOALSCLR, LEADRSHP, and REQGRAD did not have the association with achievement that was predicted in the literature on school effectiveness. Our findings suggest that having clearer goals, stronger leadership within a school, and higher expectations for students does not account for differences in average achievement among schools.

The standardized variable for CLIMACAD was strongly associated with differences in achievement in all four subjects. Our findings indicate that for every one standard deviation increase in the extent to which maintaining an academic environment is a problem in a school, estimated average student achievement decreases by between 1.79 and 2.58 points. Poor school climates appear to have the greatest impact on average achievement in mathematics and science; average reading achievement appears to be slightly less affected.

TABLE 3: Effects of School Context and Policy Characteristics on Predictors of Achievement in Reading, Mathematics, Science, and History using MEANS models

Conditional Models		Reading		Math		Science		History	
Fixed Effects									
Student Level	School Level	coeff	p	coeff	p	coeff	p	coeff	p
Intercept B0									
	G00 intercept	***49.37***	***.000***	***49.91***	***.000***	***49.01***	***.000***	***49.44***	***.000***
	G01 zclimacad	***-1.79***	***.000***	***-2.58***	***.000***	***-2.43***	***.000***	***-2.00***	***.000***
	G02 zgoalscl	-0.36	.302	-0.43	.260	-0.42	.270	-0.43	.302
	G03 zleadrsh	0.16	.637	-0.05	.897	-0.01	.970	0.37	.356
	G04 zpctfarm	***-0.84***	***.003***	***-0.94***	***.002***	***-0.92***	***.003***	***-0.70***	***.035***
	G05 zpubsch	-0.13	.692	0.10	.782	0.40	.262	0.13	.738
	G06 zreqgrad	-0.25	.289	-0.16	.544	-0.07	.797	-0.04	.886
	G07 ztquality	-0.52	.052	***-0.76***	***.010***	***-0.77***	***.008***	***-0.65***	***.043***
SES B1									
	G10 intercept	***1.30***	***.000***	***1.20***	***.000***	***1.48***	***.000***	***1.71***	***.000***
	G11 zclimacad	-0.12	.650	***-0.62***	***.013***	***-0.58***	***.033***	-0.22	.448
	G12 zgoalscl	0.19	.477	0.10	.697	0.12	.661	0.02	.953
	G13 zleadrsh	-.45	.074	-0.28	.232	-.33	.220	-0.11	.699
	G14 zpctfarm	0.09	.672	-0.24	.204	-.00	.993	-0.04	.846
	G15 zpubsch	-0.01	.965	0.50	.053	0.40	.163	0.31	.314
	G16 zreqgrad	-0.08	.645	0.15	.376	0.09	.628	0.11	.587
	G17 ztquality	.0.16	.456	0.04	.847	0.07	.736	0.07	.757
EXPECT B2									
	G20 intercept	***1.57***	***.000***	***1.62***	***.000***	***1.21***	***.000***	***1.39***	***.000***
	G21 zclimacad	-0.03	.909	0.01	.959	-0.05	.846	0.02	.944
	G22 zgoalscl	-0.15	.591	0.16	.828	-0.17	.504	-0.10	.704
	G23 zleadrsh	0.13	.618	0.08	.744	0.20	.418	-0.00	.985
	G24 zpctfarm	-0.36	.075	-0.19	.356	-0.06	.764	-0.29	.153
	G25 zpubsch	0.04	.881	-0.19	.501	-0.08	.753	-0.13	.650
	G26 zreqgrad	-0.06	.754	-0.10	.577	-0.18	.325	-0.06	.727
	G27 ztquality	0.14	.526	-0.12	.588	-0.02	.930	-0.04	.862
PERFORM B3									
	G30 intercept	***2.33***	***.000***	***3.08***	***.000***	***3.09***	***.000***	***2.60***	***.000***
	G31zclimacad	-0.46	.071	-0.43	.110	***-0.64***	***.013***	***-0.52***	***.050***
	G32 zgoalscl	0.28	.274	-0.09	.751	0.18	.485	0.47	.079
	G33 zleadrsh	-0.10	.669	0.03	.920	-0.24	.325	-0.29	.253
	G34 zpctfarm	-0.18	.376	-0.00	.985	-0.20	.324	-0.10	.656
	G35 zpubsch	-0.30	.249	-0.20	.454	-0.36	.159	-0.23	.401
	G36 zreqgrad	-0.00	.993	-0.10	.591	0.28	.101	-0.12	.496
	G37 ztquality	-0.12	.571	-0.13	.559	-0.24	.238	-0.22	.304
Variance Components (df = 233)									
		Var.	p	Var.	p	Var.	p	Var.	p
	Tau U0	8.26	.000	10.75	.000	10.25	.000	12.92	.000
	Tau U1	2.27	.000	1.69	.000	2.83	.000	3.28	.000
	Tau U2	3.17	.000	3.04	.000	2.53	.000	2.61	.000
	Tau U3	2.58	.000	3.44	.000	2.90	.000	2.83	.000
σ^2 R		25.79		24.80		25.95		28.64	

One finding at odds with our expectations was the effect of the standardized variable for TQUALITY on student achievement. Previous studies indicated that high levels of teacher accountability were associated with more desirable average school achievement. Our findings suggest that this variable, when significant, is associated with a decrease in average school achievement. This variable appeared to have the least effect in reading, (coeff = 0.52). Further study is necessary to determine what student-teacher interactions might account for this discrepancy.

Overall, none of the variables we identified were consistently useful for explaining variance in the slope coefficients for three within-school parameters. CLIMACAD did explain a significant amount of the variance in the average effect of SES on mathematics and on science achievement scores. However, the reliability of this parameter was very low, and follow-up analysis suggested that the proportion of total variance around SES explained by the school effects variables was quite low.

The total variance of each equation is composed of sampling error and parameter variance. Reliability is the proportion of total variance around each parameter that is parameter variance, and thus available to be explained. HLM calculates the reliability for each parameter using methods described by Bryk and Raudenbush (1992). Typically, the intercept is the most reliable parameter. In our models, the reliability for the intercept ranged from .526 for reading to .586 for history. Reliabilities of the slope coefficients were considerably less. All reliabilities are presented in Table 4.

Our conditional models used seven school effects variables to explain variance around each parameter. R^2 is the proportion of parameter variance explained by the model. The researcher makes this calculation by comparing the amount of variance around each parameter in the unconditional model with the remaining unexplained variance around each parameter in the conditional model. The R^2 values for each parameter are presented in Table 4.

TABLE 4: HLM Statistics for the MEANS models

Conditional Models	Reading		Math		Science		History	
Reliability (R_{xx})	B0	.526	B0	.579	B0	.563	B0	.586
(Parameter Var/ Total Variance)	B1	.278	B1	.239	B1	.313	B1	.322
	B2	.369	B2	.369	B2	.330	B2	.319
	B3	.342	B3	.400	B3	.342	B3	.340
Proportions Parameter Variance	Tau U0	.29	Tau U0	.37	Tau U0	.35	Tau U0	.20
Explained (R^2)	Tau U1	.05	Tau U1	.11	Tau U1	.03	Tau U1	.00
$(V_u - V_c) / V_u$	Tau U2	.01	Tau U2	.00	Tau U2	.00	Tau U2	.00
	Tau U3	.12	Tau U3	.02	Tau U3	.11	Tau U3	.11
Total Variance Explained by Model	Tau U0	.15	Tau U0	.22	Tau U0	.20	Tau U0	.12
(R_{xx} * R^2)	Tau U1	.01	Tau U1	.03	Tau U1	.01	Tau U1	.00
	Tau U2	.00	Tau U2	.00	Tau U2	.00	Tau U2	.00
	Tau U3	.04	Tau U3	.00	Tau U3	.04	Tau U3	.04

The total variance explained by the model is the product of reliability times R^2. The results of this model show that the variables selected for this analysis predicted changes in average achievement (the intercept) better than they predicted the effects of SES, EXPECT, and PERFORM (the slope coefficients). The four models explained between 12% (history) and 22% (mathematics) of the variance in achievement. However, low reliabilities for the slope coefficients coupled with low R^2 values suggest that the school effects in our models were not particularly useful for explaining why the strength and direction of associations of achievement with SES, EXPECT, and PERFORM vary for children in different schools. There may be other, untested variables that would provide some explanation of this variance.

These findings support previous studies that indicate that there may be more free variance to account for in mathematics than in reading. The proportion of variance in mathematics achievement explained by this model was 22%, compared to 15% for reading achievement. A similar trend was observed for science and history. The association of school effectiveness indicators with student achievement may be lower when assessments rely heavily on reading and writing skills. Estimates of associations of school effects with achievement may be more encouraging if inferences rely on data from a range of assessments.

Consequences of Replacing Missing Values with Means

We cross-validated our first analysis using a model in which cases with missing values on any of the three within-school or seven between-school variables were deleted (Listwise Deletions model). This reduced our sample from 7462 students in 247 schools to a sample of 3449 students in 136 schools. HLM deleted five more of these schools from the analysis because the number of cases per school was low or because there was no variance within the school for a variable. Our findings are based on students in 131 schools. The conditional model for this series of analyses used the same between school equations as before, and all variables were standardized to have a mean of 0 and a standard deviation of 1.

The unconditional models using listwise deletions are listed in Table 5. Although the intercept and the average values for the coefficients are slightly higher, they do not appear to be significantly different from the MEANS model. However, the amount of variance around each parameter is higher, and the residual variance is approximately twice as high as with the MEANS model. For example, the unconditional σ^2 for reading is 35.94 for the MEANS model, but 67.77 for the Listwise Deletions model. A similar trend occurs across all subject areas. One explanation for this may be that the variance tends to be underestimated when missing values are replaced with means. However, the proportion of variance explained by the three level-one predictors is very close in the two models. Although the data is not presented here, the three student level variables explained approximately 29% the within-school variance for each subject.

As expected, the intercept is the greatest predictor of achievement. The order of association for the within-school variables, PERFORM, SES, and EXPECT is consistent with the order demonstrated by variables on the MEANS models for science and history.

The pattern of the associations is robust across subject areas and is consistent with the pattern observed in the MEANS model. One standard deviation increase in PERFORM was associated with a 2.97 to 3.77 increase in achievement. As with the MEANS model, the association of PERFORM with achievement was strongest in mathematics. The patterns of association of EXPECT with achievement in mathematics and reading were the same with the Listwise Deletions model as with the MEANS model. One standard deviation increase in EXPECT predicted an increase in achievement of 1.14 points in mathematics, and 1.05 points in reading. SES was most strongly associated with achievement in history (2.17), then science (2.09), then reading (1.97), then mathematics (1.82). Perhaps differences in achievement in "basic" subjects like reading and mathematics are more weakly associated with SES differences because schools are more sensitive to overcoming the gap in these areas. If so, our study did not explain what policies reflect this sensitivity.

TABLE 5: Comparison of Average Within-School Predictors of Achievement in Reading, Mathematics, Science, and History using Listwise Deletions model

Unconditional Models		**Reading**		**Math**		**Science**		**History**	
Fixed Effects									
Student Level	School Level	Coeff	p	Coeff	p	Coeff	p	Coeff	p
Intercept B0									
	G00 intercept	51.19	.000	51.59	.000	50.40	.000	50.83	.000
SES B1									
	G10 intercept	1.97	.000	1.82	.000	2.09	.000	2.17	.000
EXPECT B2									
	G20 intercept	1.05	.000	1.14	.000	0.96	.000	0.66	.009
PERFORM B3									
	G30 intercept	2.97	.000	3.77	.000	3.39	.000	3.56	.000
Variance Components (df = 130)									
		Var.	p	Var.	p	Var.	p	Var.	p
	Tau U0	11.80	.000	21.24	.000	19.11	.000	20.08	.000
	Tau U1	5.34	.000	6.52	.000	6.29	.000	7.39	.000
	Tau U2	3.27	.000	4.00	.000	2.59	.000	3.15	.000
	Tau U3	3.22	.000	2.86	.000	3.55	.000	4.06	.000
σ^2 R		48.16		44.13		50.88		56.50	

As before, the conditional models used seven school effects to try to explain the variation around each student level parameter. The results of these tests are presented in Table 6. None of these variables explained a significant amount of variance in the slope coefficients except for the association of TQUALITY with the average effect of SES on mathematics achievement. Overall, the effect of TQUALITY, though not significant, was at least suggestive of a positive association with the mean effect of SES and with the mean effect of EXPECT. TQUALITY had a negative, although insignificant, association with PERFORM, however, which supported the results of analyses of the MEANS models.

TABLE 6: Effects of School Context and Policy Characteristics on Predictors of Achievement in Reading, Mathematics, Science, and History using Listwise Deletions models									
Conditional Models		**Reading**		**Math**		**Science**		**History**	
Fixed Effects									
Student Level	School Level								
Intercept B0									
	G00 intercept	***51.26***	***.000***	***51.73***	***.000***	***50.34***	***.000***	***50.93***	***.000***
	G01 zclimacad	***2.42***	***.000***	***-3.46***	***.000***	***-3.45***	***.000***	***-3.34***	***.000***
	G02 zgoalscl	-0.80	.089	***-1.05***	***.036***	***-1.06***	***.040***	-0.67	.246
	G03 zleadrsh	0.49	.256	0.48	.324	0.47	.351	0.71	.205
	G04 zpctfarm	-0.69	.154	***-1.63***	***.002***	***-1.41***	***.009***	-0.82	.167
	G05 zpubsch	-0.41	.396	-0.87	.089	***-1.53***	***.004***	***-1.24***	***.035***
	G06 zreqgrad	-0.09	.779	-0.25	.468	0.35	.326	0.22	.570
	G07 ztquality	***-1.08***	***.010***	***-1.01***	***.021***	***-0.90***	***.048***	-0.98	.054
SES B1									
	G10 intercept	***1.59***	***.000***	***1.30***	***.000***	***1.82***	***.000***	***1.71***	***.000***
	G11 climacad	.513	.359	-0.68	.233	0.08	.893	0.54	.400
	G12 zgoalscl	.378	.505	0.07	.886	0.11	.826	0.33	.561
	G13 zleadrsh	-0.91	.064	-0.59	.234	-0.20	.688	-0.55	.325
	G14 zpctfarm	-0.46	.353	-0.38	.446	-0.33	.517	-0.46	.410
	G15 zpubsch	0.22	.665	-1.02	.044	0.11	.830	-0.26	.654
	G16 zreqgrad	-0.15	.637	0.00	.998	0.11	.740	0.27	.452
	G17 ztquality	0.64	.136	***1.00***	***.019***	0.78	.071	.0.76	.115
EXPECT B2									
	G20 intercept	***1.01***	***.001***	***1.11***	***.000***	***0.89***	***.002***	***0.61***	***.047***
	G21 zclimacad	-0.26	.585	-0.16	.754	-0.33	.472	0.12	.811
	G22 zgoalscl	0.20	.606	0.18	.653	0.09	.806	0.17	.677
	G23 zleadrsh	0.04	.916	0.25	.519	0.19	.606	-0.03	.932
	G24 zpctfarm	0.40	.274	0.16	.683	0.33	.338	0.11	.775
	G25 zpubsch	-0.38	.384	-0.34	.446	-0.32	.451	-0.08	.871
	G26 zreqgrad	-0.13	.618	0.25	.334	-0.19	.424	-0.04	.873
	G27 ztquality	0.31	.367	0.08	.836	0.00	.993	0.28	.444
PERFORM B3									
	G30 intercept	***2.92***	***.000***	***3.77***	***.000***	***3.48***	***.000***	***3.63***	***.000***
	G31zclimacad	-0.08	.868	-0.26	.558	-0.82	.080	-0.95	.057
	G32 zgoalscl	0.68	.082	-0.13	.737	0.13	.752	0.41	.337
	G33 zleadrsh	-0.56	.132	-0.05	.892	-0.37	.330	-0.52	.555
	G34 zpctfarm	-0.55	.156	-0.36	.326	-0.49	.215	-0.25	.555
	G35 zpubsch	0.22	.568	0.15	.686	-0.00	.994	-0.03	.940
	G36 zreqgrad	-0.05	.838	-0.23	.309	0.24	.313	-0.02	.945
	G37 ztquality	-0.10	.774	-0.22	.507	-0.14	.684	-0.26	.489
Variance Components (df = 130)									
		Var.	p	Var.	p	Var.	p	Var.	p
	Tau U0	7.64	.000	10.11	.000	10.29	.000	13.41	.000
	Tau U1	5.21	.000	5.98	.000	6.04	.000	7.30	.000
	Tau U2	3.24	.000	4.25	.000	2.63	.000	3.33	.000
	Tau U3	1.75	.000	2.74	.000	3.11	.000	3.46	.000
σ^2 R		48.20		44.10		50.89		56.61	

The associations of school context variables with achievement varied somewhat when cases with missing values were deleted from the analysis. Unlike the MEANS model, the standardized school context variable for PCNTFARM explained a significant amount of variance in average achievement only in mathematics and science, not in all subjects. On average, student achievement was expected to decrease by 1.63 points in mathematics and 1.41 points in reading for every standard deviation increase in the number of students in the school who received free or reduced meals. Another difference from the MEANS model is that when this school context variable was controlled, school sector explained a significant amount of the variation in average school achievement in science and history. On average, attending public school is associated with lower achievement scores. The Listwise Deletions models estimated that the gap between the scores of public school and non-public school students was 1.53 points in science and 1.24 points in history.

The Listwise Deletions model confirms results of previous analysis with MEANS models that suggest that having clear goals, strong leadership within a school, and high expectations for students does not account for differences in average achievement among schools. As with the MEANS models, the CLIMACAD was strongly associated with differences in achievement in all four subjects. Our findings indicate that for every standard deviation increase in the extent to which maintaining an academic environment is a problem in a school, average student achievement was expected to decrease by between 2.42 and 3.46 points. The pattern of the strength of this association with achievement in different subject areas is the same as with the MEANS models. When school climate issues are more of a problem, the effects are associated more strongly with a decrease in average achievement in mathematics, science, and history. Average reading achievement within a school appears to be less affected by problems associated with poor climate than is achievement in other areas.

The models demonstrate some differences with regard to how well the level-2 variables predict the intercept. Overall, CLIMACAD and PCNTFARM are the most consistent predictors of average student achievement across subject areas. TQUALITY has an unexpected significant negative association with student achievement in both models across most subject areas. One explanation for this may be that teachers who report a greater frequency of adjusting instruction and curriculum also work with targeted groups of students who are less likely to succeed or are otherwise at risk. Further analysis of the data using other models may explain this incongruity.

HLM Statistics for the Listwise Deletions model are presented in Table 7. The reliabilities for all parameters were virtually the same across models and subject areas. This means that the models were equivalent with regard to the proportion of variance for each parameter that could have been explained by the level-2 variables. Thus, most of the difference in the total variance explained by each model appears to be associated with how well level-2 variables explain variance in the intercept. Overall, these variables were able to explain more of the variance in student achievement across subject areas when cases were deleted listwise, not when missing values were replaced with means. The Listwise Deletions models explained 4% more of the variance in average reading achievement, 9% more of the variance in average math achievement, 6% more of the variance in average science achievement, and 8% more of the variance in average history achievement between schools.

TABLE 7: HLM Statistics for the Listwise Deletions models								
Conditional Models	**Reading**		**Math**		**Science**		**History**	
Reliability (R_{xx})	B0	.526	B0	.591	B0	.569	B0	.596
(Parameter Var/ Total Variance)	B1	.352	B1	.393	B1	.368	B1	.384
	B2	.385	B2	.449	B2	.339	B2	.362
	B3	.374	B3	.371	B3	.368	B3	.368
Proportions Parameter Variance	Tau U0	.35	Tau U0	.52	Tau U0	.46	Tau U0	.33
Explained (R^2)	Tau U1	.02	Tau U1	.08	Tau U1	.04	Tau U1	.01
$(V_u - V_c) / V_u$	Tau U2	.01	Tau U2	.00	Tau U2	.00	Tau U2	.00
	Tau U3	.46	Tau U3	.04	Tau U3	.12	Tau U3	.15
Total Variance Explained by Model	Tau U0	.19	Tau U0	.31	Tau U0	.26	Tau U0	.20
(R_{xx} * R^2)	Tau U1	.01	Tau U1	.03	Tau U1	.01	Tau U1	.00
	Tau U2	.00	Tau U2	.00	Tau U2	.00	Tau U2	.00
	Tau U3	.17	Tau U3	.02	Tau U3	.05	Tau U3	.05

Consequences of Redefining of School Climate

Five different descriptions of school climate emerged from our principal components analysis of the data. The purpose of our third group of comparisons was to estimate how substitution of these five alternative school climate variables affected interpretation of the unique association of school climate with mean school achievement in four subject areas. In each model, the association of school climate was estimated after six other school effects were controlled. The slope coefficients were allowed to vary randomly. Using random slope coefficients allowed the average effect of each level-one variable to be adjusted to represent the average effect of that variable on all students within the school, but no attempt was made to explain variation in the strength and direction of the associations of variables with achievement from one school to another. Support for our decision to use a more parsimonious model for the climate variable comparisons came from our first analysis. Although the slope coefficients vary significantly, the school effects we selected explained a negligible proportion of total variance around each slope coefficient.

The unconditional models for each subject area were the same as for those presented in Table 2. The amount of variance in the intercept that could be predicted by school effects was 11.65 in reading, 17.19 in mathematics, 15.87 in science, and 16.22 in history. Conditional models comparing the effects of changes in the definition of school climate on average achievement in reading are presented in Table 8. Consequences of respecification of the construct school climate in other subject areas are presented in Tables 9 (mathematics), 10 (science) and 11 (history).

The conditional models for reading appear robust with regard to changes in the definitions of the construct school climate. In every model, school climate and the percent of students receiving free and reduced meals (PCNTFARM) are significant predictors of school achievement when other variables are controlled. In two models, teacher accountability emerged as a third significant predictor, although the direction of this association is the reverse of what was hypothesized in the literature. TQUALITY significantly reduces estimates of student achievement when school climate is defined in terms of student behavior (CLIMBEHV) or the extent to which violence and crime are problems within the school (CLIMVICR). Further investigation of this interaction might be helpful for explaining why teacher expectations and student achievement are negatively correlated in our study.

TABLE 8: Effects of Respecification of School Climate Characteristics on Predictors of Average Achievement in Reading

Conditional Models	**climacad**		**climatnd**		**climbehv**		**climrlen**		**climvicr**	
	coeff	p	coeff	p	coeff	p	coeff	p	coeff	p
Fixed Effects										
Intercept B0										
G00 intercept	***49.48***	***.000***	***49.44***	***.000***	***49.51***	***.000***	***49.43***	***.000***	***49.48***	***.000***
G01 zclimate	***-1.62***	***.000***	***-1.34***	***.000***	***-1.62***	***.000***	***0.87***	***.006***	***-1.03***	***.001***
G02 zgoalscl	-0.49	.143	-0.51	.129	-0.44	.181	-0.38	.261	-0.43	.2115
G03 zleadrsh	0.30	.344	0.20	.528	-0.04	.899	-0.13	.712	0.22	.505
G04 zpctfarm	***-0.67***	***.009***	***-0.81***	***.002***	***-0.76***	***.002***	***-1.09***	***.000***	***-0.83***	***.002***
G05 zpubsch	-0.19	.559	-0.25	.448	-0.50	.093	***-0.82***	***.006***	-0.59	.058
G06 zreqgrad	-0.22	.340	-0.20	.388	-0.23	.313	-0.21	.367	-0.29	.217
G07 ztquality	-0.46	.077	-0.29	.272	***-0.67***	***.012***	-0.36	.175	-0.29	.279
SES B1										
G10 intercept	***1.32***	***.000***	***1.33***	***.000***	***1.32***	***.000***	***1.34***	***.000***	***1.32***	***.000***
EXPECT B2										
G20 intercept	***1.54***	***.000***	***1.57***	***.000***	***1.54***	***.000***	***1.58***	***.000***	***1.55***	***.000***
PERFORM B3										
G30 intercept	***2.22***	***.000***	***2.23***	***.000***	***2.21***	***.000***	***2.22***	***.000***	***2.22***	***.000***
Variance Components (Tau UO df = 233; Tau U1 - Tau U3 df = 240)										
	Var.	p	Var.	p	Var.	p	Var.	p	Var.	p
Tau U0	8.34	.000	8.58	.000	8.42	.000	9.20	.000	8.87	.000
Tau U1	2.33	.000	2.33	.000	2.34	.000	2.36	.000	2.34	.000
Tau U2	3.08	.000	3.11	.000	3.06	.000	3.11	.000	3.09	.000
Tau U3	2.83	.000	2.82	.000	2.82	.000	2.82	.000	2.82	.000
σ^2 R	25.79	.000	25.80		25.80		25.81		25.80	

The conditional models for mathematics and science achievement, shown in Tables 9 and 10, are quite comparable. Regardless of the definition of school climate used in this study, school climate (CLIM*) and percent of students who receive free and reduced meals (PCNTFARM) are significant predictors of student achievement when other variables are controlled. As with the reading achievement models, teacher accountability (TQUALITY) is negatively associated with achievement when teacher reports regarding student behavior are the basis of establishing a definition of school climate. Unlike reading achievement, mean mathematics and science achievement scores were negatively associated with TQUALITY when climate was defined as academic environment.

Whether or not students attend public school is significantly associated with reading and mathematics achievement when the single item variable CLIMRLEN is used as a variable. However, the gap in mean achievement between students who attend public or private schools is insignificant when composite definitions of school climate are substituted. Perhaps problems identified in the composite variables, such as poor attendance, student misbehavior, and violence and crime, are more prevalent on average in public schools, and have the greatest association with basic skill achievement. Once these problems are identified as school climate issues, or when alternative achievement measures are considered, the association of school sector with achievement may be inconsequential.

TABLE 9: Effects of Respecification of School Climate Characteristics on Predictors of Average Achievement in Mathematics

Conditional Models	**climacad**		**climatnd**		**climbehv**		**climrlen**		**climvicr**	
	coeff	p	coeff	p	coeff	p	coeff	p	coeff	p
Fixed Effects										
Intercept B0										
G00 intercept	***50.07***	***.000***	***50.02***	***.000***	***50.10***	***.000***	***50.01***	***.000***	***50.07***	***.000***
G01 zclimate	***-2.23***	***.000***	***-1.78***	***.000***	***-2.05***	***.000***	***0.75***	***.028***	***-1.54***	***.000***
G02 zgoalscl	-0.51	.149	-0.53	.143	-0.40	.250	-0.25	.489	-0.47	.197
G03 zleadrsh	0.12	.726	-0.03	.941	-0.21	.536	-0.28	.461	0.01	.983
G04 zpctfarm	***-0.72***	***.008***	***-0.89***	***.001***	***-0.87***	***.001***	***-1.24***	***.000***	***-0.87***	***.002***
G05 zpubsch	-0.03	.944	-0.15	.683	-0.54	.097	***-1.01***	***.002***	-0.53	.109
G06 zreqgrad	-0.14	.573	-0.11	.647	-0.17	.475	-0.16	.518	-0.25	.324
G07 ztquality	***-0.66***	***.017***	-0.43	.123	***-0.91***	***.002***	-0.50	.080	-0.42	.138
SES B1										
G10 intercept	***1.22***	***.000***	***1.23***	***.000***	***1.22***	***.000***	***1.25***	***.000***	***1.22***	***.000***
EXPECT B2										
G20 intercept	***1.57***	***.000***	***1.59***	***.000***	***1.56***	***.000***	***1.60***	***.000***	***1.58***	***.000***
PERFORM B3										
G30 intercept	***3.04***	***.000***	***3.05***	***.000***	***3.02***	***.000***	***3.02***	***.000***	***3.03***	***.000***
Variance Components (Tau UO df = 233; Tau U1 - Tau U3 df = 240)										
	Var.	p	Var.	p	Var.	p	Var.	p	Var.	p
Tau U0	10.98	.000	11.69	.000	11.20	.000	13.19	.000	11.78	.000
Tau U1	1.97	.000	2.03	.000	1.95	.000	2.044	.000	2.00	.000
Tau U2	2.92	.000	2.97	.000	2.86	..000	2.93	.000	2.92	.000
Tau U3	3.39	.000	3.42	.000	3.41	.000	3.46	.000	3.43	.000
σ^2 R	24.79		24.79		24.82		24.81		24.79	

The comparability of our hierarchical models for evaluating the association of school effectiveness indicators and student achievement in mathematics and science suggests that the kinds of cognitive skills measured by tests in these two areas are similar. If so, either achievement test may be sufficient for estimating the associations of school effects with cognitive skills that involve quantitative reasoning. Concerns that a broad range of tests are necessary to evaluate the relationship of school effects with student achievement may be addressed by using assessments that tap a broad range of outcomes.

TABLE 10: Effects of Respecification of School Climate Characteristics on Predictors of Average Achievement in Science

Conditional Models	**climacad**		**climatnd**		**climbehv**		**climrlen**		**climvicr**	
	coeff	p	coeff	p	coeff	p	coeff	p	coeff	p
Fixed Effects										
Intercept B0										
G00 intercept	***49.13***	***.000***	***49.08***	***.000***	***49.19***	***.000***	***49.06***	***.000***	***49.14***	***.000***
G01 zclimate	***-2.03***	***.000***	***-1.81***	***.00***	***-1.82***	***.000***	***1.08***	***.002***	***-1.60***	***.000***
G02 zgoalscl	-0.55	.122	-0.61	.090	-0.52	.135	-0.40	.264	-0.54	.133
G03 zleadrsh	0.23	.494	0.10	.779	0.04	.914	-0.29	.434	0.11	.755
G04 zpctfarm	***-0.80***	***.004***	***-0.933***	***.001***	***-0.93***	***.001***	***-1.30***	***.000***	***-0.91***	***.001***
G05 zpubsch	0.31	.386	0.29	.410	-0.14	.651	-0.48	.131	-0.04	.886
G06 zreqgrad	0.15	.534	-0.12	.612	-0.24	.328	-0.18	.467	-0.25	.314
G07 ztquality	***-0.64***	***.021***	-0.44	.116	***-0.78***	***.006***	-0.53	.061	-0.41	.140
SES B1										
G10 intercept	***1.53***	***.000***	***1.55***	***.000***	***1.42***	***.000***	***1.56***	***.000***	***1.53***	***.000***
EXPECT B2										
G20 intercept	***1.17***	***.000***	***1.19***	***.000***	***1.19***	***.000***	***1.21***	***.000***	***1.18***	***.000***
PERFORM B3										
G30 intercept	***2.91***	***.000***	***2.92***	***.000***	***2.93***	***.000***	***2.91***	***.000***	***2.91***	***.000***
Variance Components (Tau UO df = 233; Tau U1 - Tau U3 df = 240)										
	Var.	p	Var.	p	Var.	p	Var.	p	Var.	p
Tau U0	10.68	.000	10.92	.000	11.15	.000	12.28	.000	10.93	.000
Tau U1	2.94	.000	2.99	.000	0.35	.000	2.99	.000	2.92	.000
Tau U2	2.44	.000	2.48	.000	2.25	.000	2.47	.000	2.45	.000
Tau U3	3.07	.000	3.10	.000	3.09	.000	3.17	.000	3.14	.000
σ^2 R	25.95		25.94		26.56		25.96		25.95	

The effects of respecification of school climate variables on history achievement is shown in Table 11. The amount of parameter variance in mean achievement among schools was greater in history than in science or in mathematics, but this variance was less associated with the school effects variables we selected than was variance in other subject areas. Although some combination of school effects explained more than thirty percent of the differences in mean achievement in mathematics and science, and as much as 28% of the variation in reading, only 14% to 20% of the variation in history achievement among schools was explained by a combination of school climate and other school effects. The patterns of associations were consistent, overall, with those for other subject areas: school climate and percent of students in the school who receive free and reduced meals were significant predictors of mean history achievement when other variables were controlled. One exception was that PCNTFARM was not a significant predictor of history achievement when academic environment (CLIMACAD) was controlled. As with other models, TQUALITY was negatively associated with mean history achievement, particularly when student behavior (CLIMBEHV) was reported by teachers to be a problem in their schools. Yet despite the consistency of the pattern of significance of school effects, and the potential for explanation of this parameter, our models were weakest for predicting achievement in history. Variables other than those used in this study may be useful for identifying relationships between school effects and history achievement.

TABLE 11: Effects of Respecification of School Climate Characteristics on Predictors of Average Achievement in History

Conditional Models	**climacad**		**climatnd**		**climbehv**		**climrlen**		**climvicr**	
	coeff	p	coeff	p	coeff	p	coeff	p	coeff	p
Fixed Effects										
Intercept B0										
G00 intercept	***49.53***	***.000***	***49.50***	***.000***	***49.55***	***.000***	***49.49***	***.000***	***49.53***	***.000***
G01 zclimate	***-1.79***	***.000***	***-1.57***	***.000***	***-1.67***	***.000***	***0.92***	***.015***	***-1.20***	***.001***
G02 zgoalscl	-0.63	.115	-0.68	.092	-0.56	.155	-0.42	.200	-0.59	.151
G03 zleadrsh	0.59	.126	0.48	.223	0.32	.399	0.14	.732	0.51	.196
G04 zpctfarm	-0.54	.076	***-0.67***	***.026***	***-0.67***	***.024***	***-1.01***	***.001***	***-0.69***	***.024***
G05 zpubsch	-0.08	.838	0.04	.911	-0.29	.410	-0.64	.065	-0.33	.369
G06 zreqgrad	0.00	.997	0.03	.927	-0.01	.969	0.14	.959	-0.07	.797
G07 ztquality	-0.55	.076	-0.37	.230	***-0.74***	***.020***	-0.44	.165	-0.36	.254
SES B1										
G10 intercept	***1.77***	***.000***	***1.78***	***.000***	***1.77***	***.000***	***1.80***	***.000***	***1.78***	***.000***
EXPECT B2										
G20 intercept	***1.34***	***.000***	***1.36***	***.000***	***1.34***	***.000***	***1.37***	***.000***	***1.35***	***.000***
PERFORM B3										
G30 intercept	***2.53***	***.000***	***2.54***	***.000***	***2.52***	***.000***	***2.52***	***.000***	***2.52***	***.000***
Variance Components (Tau UO df = 233; Tau U1 - Tau U3 df = 240)										
	Var.	p	Var.	p	Var.	p	Var.	p	Var.	p
Tau U0	12.98	.000	13.14	.000	13.14	.000	13.95	.000	13.51	.000
Tau U1	3.12	.000	3.13	.000	3.13	.000	3.18	.000	3.11	.000
Tau U2	2.54	.000	2.55	.000	2.52	.000	2.54	.000	2.54	.000
Tau U3	3.09	.000	3.09	.000	3.07	.000	3.11	.000	3.09	.000
σ^2 R	28.63		28.64		28.65		28.65		28.64	

Table 12 reports the relative performance of different definitions of school climate variables as predictors of achievement when six other school effects are controlled. The amount of total variance in achievement in four subject areas explained by different climate variables varied considerably. The model using academic environment (CLIMACAD) as a predictor of mathematics achievement predicted 21% of the variance in average achievement among schools. Only 8% of the variation in mean history achievement among schools was predicted by using a definition of school climate based on the extent to which teachers reported that violence and crime was a problem in their school (CLIMVICR). In general, all five models were better predictors of mean mathematics and science achievement than they were of average achievement in reading and history.

The single-item climate variable measuring school discipline (CLIMRLEN) was a significant predictor of mean school achievement, but performed more poorly in general in our models than variables based on item composites. As a predictor of history achievement, however, CLIMRLEN performed better than CLIMVICR, and virtually the same as CLIMACAD, CLIMATND, and CLIMBEHV. Subject difference were also evident. When any of the five school climate definitions was compared across subject areas, climate was consistently associated with the greatest difference in mathematics achievement, followed by achievement in science, reading and history.

TABLE 12: HLM Statistics for the Comparisons of Variance in School Mean Achievement Explained by Different Climate Variables in Different Subject Areas

Conditional Models		Reading		Math		Science		History	
Statistic	*Climate Variable*								
Reliability (R_{xx})	***climacad***	B0	.528	B0	.583	B0	.570	B0	.587
(Variance of B0/	***climatnd***	B0	.533	B0	.594	B0	.574	B0	.589
Total Variance)	***climbehv***	B0	.530	B0	.586	B0	.574	B0	.589
	climrlen	B0	.545	B0	.614	B0	.594	B0	.599
	climvicr	B0	.539	B0	.595	B0	.574	B0	.594
Proportions	***climacad***	U0	.28	U0	.36	U0	.33	U0	.20
Parameter	***climatnd***	U0	.26	U0	.32	U0	.31	U0	.19
Variance	***climbehv***	U0	.28	U0	.35	U0	.30	U0	.19
Explained (R^2)	***climrlen***	U0	.21	U0	.23	U0	.23	U0	.14
$(V_u - V_c) / V_u$	***climvicr***	U0	.24	U0	.31	U0	.31	U0	.17
	climacad	U0	.15	U0	.21	U0	.19	U0	.12
Total Variance	***climatnd***	U0	.14	U0	.19	U0	.18	U0	.12
Explained by	***climbehv***	U0	.15	U0	.20	U0	.17	U0	.11
Model	***climrlen***	U0	.11	U0	.14	U0	.13	U0	.11
(R_{xx} * R^2)	***climvicr***	U0	.13	U0	.19	U0	.18	U0	.08

CONCLUSION

School effectiveness indices appear to be robust to changes in measures of achievement. School effects that predicted success in one subject area were good predictors of achievement in another. However the effects themselves may be less significant than the literature on school effectiveness suggests. Our first series of hierarchical linear models (MEANS models) failed to support, on balance, our assumptions about the influence of school context and policy variables for school effectiveness. Once school socioeconomic status was controlled (PCNTFARM), only our school climate variable (CLIMACAD) explained differences in mean achievement across all subject areas. And, contrary to our expectations, teacher accountability (TQUALITY) was negatively associated with mean school achievement in mathematics, science, and history. Our second series of models, the Listwise Deletions models, showed an almost identical pattern of associations of school effects with achievement across four subject areas.

The proportion of total variance explained by our models was higher with mathematics and science models than with reading and history models. Differences in the amount of total variation explained by different groups of models support claims that the amount of free variation is greater in some subjects than in others. Although most of the studies in this area have reported differences in elementary school achievement, our results suggest that this difference may generalize more broadly. Our findings indicate that in high school, the amount of free variation, that is the extent to which school effects are associated with differences in various subject areas, is greater in subjects that emphasize non-verbal, quantitative cognitive skills than those that rely heavily on verbal skills of the type required in reading or history.

The hierarchical models appear to be sensitive to differences in the way that missing data are handled. Although HLM offers missing data options at level-1, large numbers of missing values tends to reduce the number of parameters than can be estimated at level-2 and/or reduce the number of schools that are included in the analysis. Further, since HLM tolerates no missing data at level-2, schools missing data for any variable included in a model must be dropped from the analysis. One method for handling missing data is to substitute mean values for missing values, as was done with the MEANS database. An alternative is to delete cases with missing data. In this study, the Listwise Deletions database reduced the original sample size by about one-half. The strength and direction of the associations of predictors with their parameters generally was robust whether cases with missing data were deleted (Listwise Deletions models) or whether all cases were retained by replacing missing values of variables with the means of those variable (MEANS models). However, the models yielded different estimates of the proportion of total variance around each parameter that may be explained by a set of predictors. A simulation study would be useful to determine if the MEANS models underestimation of the variance in school mean achievement that can be explained by school effects is a greater problem than that found from the use of the smaller sample sizes of Listwise Deletions models.

Interpretation of the effect of school climate as a measure of school effectiveness is complicated by differences in the ways in which the construct is defined across a range of studies. In most studies, the number of items available for constructing a composite is a limiting factor in developing an operational definition of a construct. That was not the case in our study. The teacher questionnaire used to evaluate school climate in the HSES study includes a section labeled "School Climate" that contains seventy-four items grouped into eleven general questions. Principal components analysis of these items suggested that groups of items explained different facets of what might be meant by school climate, and selection of specific composites would depend on the focus of the researcher. Our findings indicate that hierarchical models may be robust to differences in the specification of the definition of this variable. A table of correlations of the standardized climate variables we selected for this study is presented in the appendix.

Determining whether implementation of specific policies makes a difference in student achievement over and above what students would be expected to achieve in the absence of those policies has serious implications for educators. Hierarchical models offer a promising method of making such a determination. However, some estimates of the impact of these policies and the importance of school context variables may be sensitive to differences in model specification. Thus, conclusions about the impact of these school effectiveness indicators are at best tentative. Guidelines for making decisions about how to handle missing data and how to measure theoretical constructs could improve our analyses about what elements of school structure and function positively influence student achievement.

Although our focus was on methodological issues, our findings suggest that widely-accepted beliefs about the strength of associations of school characteristics with student achievement need reevaluation. Because the literature is so consistent regarding what school policies and contexts are associated with student achievement, we expected all of the variables in our models to explain differences in student achievement in at least one subject area. That was not the case. We found four results that seem robust to the methodological issues we raised. They are:

1. Student level characteristics matter, whether measured at the school or student level, and these characteristics are invariant to school policy.

2. The willingness of a school to control problems that affect the learning environment, including problems associated with attendance, student discipline, violence, and crime, is strongly associated with differences in performance.

3. Schools should not expect to improve student performance by having stronger leadership, clearer goals, and stricter graduation requirements.

4. Efforts to improve student achievement by modifying instruction and by holding teachers more accountable for student outcomes are probably wasted, and may even have a negative impact on student performance.

It is possible that redefining the constructs using different items could produce a result more consistent with theoretical expectations about the value-added effects of school context variables. Also, school effects may be more significant in elementary school than in the high schools we studied. Further study would be useful for explaining these findings and guiding decisions about school policies and use of resources.

REFERENCES

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

Anderson, C. S. (1982). The search for school climate: a review of the research. *Review of Educational Research*, 52(2), 368-420.

Arnold, C. L. (1992). An introduction to hierarchical linear models. *Measurement and Evaluation in Counseling and Development*, 25, 58-90.

Averch, H. A., Carroll, S. J., Donaldson, T. S., Kiesling, H. J., and Pincus, J. (1974). How Effective Is Schooling? A Critical Review of Research. Englewood Cliffs: Educational Technology Publications.

Bryk, A. S. and Raudenbush, S. W. (1992). Hierarchical Linear Models: Applications and Data Analysis Methods. Newbury Park: Sage Publications.

Burstein, L. (1980). *The analysis of multilevel data in educational research and evaluation*. In Review of Research in Education, D. Berliner, (Editor). American Educational Research Association: Washington, D. C.

Burstein, L., Linn, R. L., and Capell, F. J. (1978). Analyzing multilevel data in the presence of heterogeneous within-class regressions. *Journal of Educational Statistics*, 3(4), 347-383.

Coleman, J. S., et al. (1966). Equality of educational opportunity. Washington, D. C.: U. S. Department of Health, Education and Welfare.

Crone, L. J., Lang, M. H., Teddlie, C., and Franklin, B. J. (1995). Achievement measures of school effectiveness: comparison of model stability across years. *Applied Measurement in Education*, 8(4), 353-365.

Dow, I. I. and Oakley, W. F. (1992). School effectiveness and leadership. *Alberta Journal of Educational Research*, 38(1), 33-47.

Gamoran, A. (1991). Schooling and achievement: additive versus interactive models. In Schools, Classroom, and Pupils. San Diego: Academic Press.

Goldstein, H. (1987). Multilevel Models in Educational and Social Research. New York: Oxford University Press.

Ingels, S. J., et al. (1994). NELS:88 Second Follow-Up Student Component Data File User's Manual. Washington, D. C.: United States Department of Education.

Jencks, E. L., et al. (1972). Inequality: A Reassessment of the Effects of Family and Schooling in America. New York: Basic Books.

Kennedy, E. (1992). A multilevel study of elementary male black students and white students. *Journal of Educational Research*, 86(2), 105-110.

Lee, V. E., et al., (1996). The Influence of School Climate on Gender Differences in the Achievement and Engagement of Young Adolescents. Washington, D. C.: American Association of University Women.

Mackenzie, D. E. (1983). Research for school improvement: an appraisal of some recent trends. *Educational Researcher*, 12(4), 5-17.

Mandeville, G. K. and Anderson, L. W. (1987). The stability of school effectiveness indices across grade levels and subject areas. *Journal of Educational Measurement*, 24(3), 203-216.

Mandeville, G. K. (1988). School effectiveness indices revisited: cross year stability. *Journal of Educational Measurement*, 25(4), 349-356.

Mendro, R. L., Webster, W. J., Bembry, K. L., and Orsak, T. H. (1995). An application of hierarchical linear modeling to determining school effectiveness. Paper presented at the American Educational Research Association Annual Meeting, San Francisco.

Miller, S. K. (1992). Changing conceptions of good schools: implications for reforming urban education. *Education and Urban Society*, 25(1), 71-84.

Orsak, T. H., et al. (1996) Empirical difficulties in using hierarchical linear models for school rankings. Paper presented at the American Educational Research Association Annual Meeting, New York.

Paterson, L. and Goldstein, H. (1991). New statistical methods for analysing social structures: an introduction to multilevel models. *British Educational Journal*, 17(4) 387-393.

Phillips, G. W. (1996). Practical applications of hierarchical linear models to district evaluations. Paper presented at the American Education Research Association Annual Meeting, New York.

Purkey, S. C. and Smith, M. S. (1983). Effective schools: a review. *Elementary School Journal*, 83(4), 427-452.

Raudenbush, S. W. and Willms, J. D. (1995). Estimation of school effects. *Journal of Educational and Behavioral Statistics*, 20(4), 307-335.

Rowan, B., Bossert, S. T., and Dwyer, D. C. (1983). Research on effective schools: a cautionary note. *Educational Researcher*, 12(4), 24-31.

Sammons, P., Nuttall, D., and Currance, P. (1993). Differential school effectiveness: results from a reanalysis of the Inner London Education Authority's junior school project data. *British Educational Research Journal*, 19(4), 381-405.

Scott, L. A., et al. (1995). High School Effectiveness Study: Data File User's Manual. Washington, D. C.: United States Department of Education.

Weber, G. (1971). Inner-city Children Can Be Taught to Read: Four Successful Schools. Washington D. C.: Council for Basic Education.

Webster, W. J., Mendro, R. L. and Almaguer, T. O. (1994). Effectiveness indices: a "value-added" approach to measuring school effect. *Studies in Educational Evaluation*, 20, 113-145.

Webster. W. J. and Mendro, R. L. (1995). Evaluation for improved school level-decision-making and productivity. *Studies in Educational Evaluation*, 21(4), 361-399.

Willms, J. D. and Raudenbush, S. W. (1989). A longitudinal hierarchical linear model for estimating school effects and their stability. *Journal of Educational Measurement*, 26(3), 209-232.

APPENDIX

STUDENT CHARACTERISTICS

Student level characteristics were selected based on recommendations made in the High School Effectiveness Study Data File User's Manual, Appendix I. Composites of these items were developed using principal components analysis and reliability scaling. Numbers in parentheses after each item indicate the minimum and maximum values for that item in the composite.

EXPECT (Sum 2 items)

DEGREE: How far in school R thinks he will get (1, 9)
SUR2GRAD: R sure to graduate from high school (1, 4)

PERFORM (Sum 2 items)

ACADPROG: Student has academic high school program (0, 1)
L1GRAD4(GPA): Mean of R's grades in English, Math, Science, and History (0, 4)

SES

S1SES: Socio-Economic Status Composite includes mother's and father's education levels and occupations, and family income. A list of items used for constructing the S1SES composite is available in Ingels, et al. (1994).

School Characteristics

School context variables were selected based on findings of previous analysis of school effectiveness, presented in Ingels, et al., pp. M-8, M-9 and the High School Effectiveness Study Data File User's Manual, Appendix I. School climate composites were derived from items on the school climate section of the teacher questionnaire. When necessary, items were reverse coded to simplify interpretation of the composites; reverse coded items are marked with an asterisk (*). Numbers in parentheses after each item indicate the minimum and maximum values for that item in the composite.

PUBSCH

School is a public school (0, 1)

PCNTFARM

Percent of students in school that receive free or reduced meals. (0, 100)

REQGRAD (Sum total for two items)

REQ7SUB: Composite variable; number of credits required in English, math, history, science, foreign language, music, and art. (9, 38.5)
POLCMTST: Minimum competency test required for graduation (0, 1)

GOALSCLR

IV1j: Goals and priorities for the school are clear.

LEADRSHP (MEAN 8 items; Strongly Disagree to Strongly Agree)

- IV1f: The principal does a GOOD* job of getting resources for this school. (1, 6)
- IV1g: The principal deals effectively with pressures from outside the school that might interfere with my teaching. (1, 6)
- IV1h: The principal sets priorities, makes plans, and sees that they are carried out. (1, 6)
- IV1o: The principal knows what kind of school he/she wants and has communicated it to the staff. (1, 6)
- IV1p: This school's administration knows the problems faced by the staff. (1, 6)
- IV2i: The principal lets staff know that is expected of them. (1, 6)
- IV2k: The principal is interested in innovation and new ideas. (1, 6)
- IV2m: The principal usually consults with staff members before he/she makes decision that affect us. (1, 6)

TQUALITY (MEAN (6 Items); teacher self-rating; Strongly Disagree to Strongly Agree)

- IV.5a: If I try really hard, I can get through even to the most difficult or unmotivated students.
- IV5b: I feel that it's part of my responsibility to keep students from dropping out of school.
- IV5c: If some students in my class aren't doing well, I feel that I should change my approach to the subject.
- IV5d: By trying a different teaching method, I can significantly affect a student's achievement.
- IV5e: There is MUCH* I can do to insure that most of my students achieve at a high level.
- IV5f: I am certain I am making a difference in the lives of my students.

CLIMACAD (MEAN (3a, 3b, 3c) + MEAN (1e, 2l, 2n); high value is serious problem)

IV3a-c: To what extent is each of these a problem?*

IV3a: tardiness* (1, 4)
IV3b: absenteeism* (1, 4)
IV3c: class cutting* (1, 4)

IV1e: The level of student misbehavior (e.g. noise, horseplay, or fighting in the halls, cafeteria, or student lounge) in this school interferes with my teaching) (1, 6)

IV2l: Rules for student behavior are consistently enforced in this school.

IV2n The attitudes and habits students bring to my class greatly reduce their chances of academic success. (1,6)

CLIMVICR (MEAN (all items); All items reverse coded; high value is serious problem)

IV3d-g, j-m: To what extent is each of these a problem?*

IV3d: physical conflicts* (1, 4)
IV3e: gang activity* (1, 4)
IV3f: robbery or theft* (1, 4)
IV3g: vandalism* (1, 4)
IV3j: possession of weapons* (1, 4)
IV3k: physical abuse of teachers* (1, 4)
IV3l: verbal abuse of teachers* (1, 4)
IV3m: racial/ethnic conflict among students* (1, 4)

CLIMRLEN (one item; high value is more consistent enforcement)

IV2l: Rules for student behavior are consistently enforced in this school.

CLIMATND (MEAN (3a, 3b, 3c); high value is serious problem)

To what extent is each of these a problem?*

IV3a: tardiness* (1, 4)
IV3b: absenteeism* (1, 4)
IV3c: class cutting* (1, 4)

CLIMBEHV (MEAN (1e, 2l, 2n); high value is serious problem)

IV1e: The level of student misbehavior (e.g. noise, horseplay, or fighting in the halls, cafeteria, or student lounge) in this school interferes with my teaching) (1, 6)

IV2n The attitudes and habits students bring to my class greatly reduce their chances of academic success. (1,6)

Correlations of Climate Variables

Correlations of School Climate Characteristics for MEANS Models (N = 247)

	climacad		climatnd		climbehv		climrlen		climvicr	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
climacad	1.00	.000								
climatnd	.9214	.000	1.00	*.000*						
climbehv	.8616	.000	.7007	.000	1.00	.000				
climrlen	-.3609	.000	-.5148	.000	-.5577	.000	1.00	.000		
climvicr	-8070	.000	.8119	.000	.6927	.000	-.4457	.000	1.00	.000

Correlations of School Climate Characteristics for Listwise Deletions Models (N = 136)

	climacad		climatnd		climbehv		climrlen		climvicr	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
climacad	1.00	.000								
climatnd	.9174	.000	1.00	.000						
climbehv	.8831	.000	.7148	.000	1.00	.000				
climrlen	-.3277	.000	-.5014	.000	-.4905	.000	1.00	.000		
climvicr	-8019	.000	.8021	.000	.7294	.000	-.4676	.000	1.00	.000

AREA 1997

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

ERIC®

TM026980

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Estimating School Value-Added Effectiveness: Consequences of Respecification of Hierarchical Linear Models	
Author(s): Clare E. VonSecker and Robert W. Lissitz	
Corporate Source: University of Maryland at College Park	Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

[X] Sample sticker to be affixed to document	Sample sticker to be affixed to document [ ]
Check here Permitting microfiche (4"x 6" film), paper copy, electronic, and optical media reproduction	**or here** Permitting reproduction in other than paper copy.
"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."	"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."
Level 1	**Level 2**

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature: Clare VonSecker	Position: GRADUATE STUDENT
Printed Name: CLARE VON SECKER	Organization: UNIVERSITY of MD
Address: 4515 Willard Ave #2104 Chevy CHASE, MD 20815	Telephone Number: (301) 652-2851
	Date: 4-8-96

OVER