International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

Three-Level Hierarchical Linear Modelling of Student Academic Achievement: Disentangling Student, Classroom, and School Effects in a Nested Experimental Design

G. Mokesh Rayalu[1] , K. Murali[2], Bandi Ramanjineyulu[3*]

Assistant Professor, Statistics & OR Division, School of Advanced Sciences, VIT University, Vellore
Assistant Professor, Sri Padmavathi College of Computer Science and Technology, Tiruchanoor, Tirupati
Senior Process Associate ,TCS , Bangalore ,India
Corresponding Author : ramanji.bandi@gmail.com

Abstract

Educational research data are inherently hierarchical in structure: students are nested within classrooms, which are in turn nested within schools, creating a three-level data architecture that violates the independence assumption of classical ordinary least squares regression and single-level analysis of variance. Applying conventional ANOVA or OLS regression to nested educational data produces biased standard errors and inflated Type I error rates, rendering statistical inferences misleading. The present study develops and applies a three-level hierarchical linear model (HLM) to investigate the relative contributions of student-level, classroom-level, and school-level factors to mathematics achievement among 1,478 secondary school students distributed across 39 classrooms in 8 schools in a semi-urban Indian district. A systematic model-building strategy—progressing from the unconditional null model through student-only, class-only, school-only, and full cross-level interaction specifications—was employed to partition variance across levels and to identify significant predictors at each hierarchical tier. The intraclass correlation coefficient from the null model (ICC = 0.257) confirmed that approximately 25.7% of total variance in mathematics achievement is attributable to between-school differences, justifying the multilevel analytical framework. Prior academic performance, student socioeconomic status, and motivation emerged as significant student-level predictors; class size and teacher experience as significant classroom-level predictors; and school mean SES and school management type as significant school-level predictors. Statistically significant cross-level interactions were identified between teacher experience and prior academic performance, and between school mean SES and individual SES, indicating that the benefits of experienced teachers and the penalty of low individual SES both depend on the school-level context. The full model reduced unexplained between-school variance by 72.2% relative to the null

model. The findings have substantive implications for educational policy, teacher allocation, and school resource distribution in the Indian secondary education context.

Keywords: hierarchical linear models, multilevel modeling, nested experimental designs, intraclass correlation, educational achievement, cross-level interactions, variance components, random effects, fixed effects, school effectiveness

1. Introduction

The analysis of data arising from hierarchical or nested structures constitutes one of the central methodological challenges of contemporary applied statistics and educational research. In the field of education, students are not sampled independently from a homogeneous population but are organized within naturally occurring institutional groups—classrooms and schools—that share common environments, resources, policies, and personnel. This clustering induces statistical dependencies among observations from the same group that fundamentally invalidate the independence assumption underlying classical ordinary least squares (OLS) regression and single-level analysis of variance (ANOVA) (Bryk & Raudenbush, 1992). When these dependencies are ignored, standard errors of regression coefficients are systematically underestimated, leading to inflated Type I error rates and overstated evidence for the statistical significance of predictor effects (Goldstein, 2003; Snijders & Bosker, 1999).

Hierarchical linear models (HLM)—also referred to as multilevel models or mixed linear models—were developed precisely to address this structural reality. By explicitly representing the nested architecture of the data through a system of linked regression equations at each hierarchical level, HLM simultaneously models within-group relationships at the individual level and between-group variation attributable to group-level characteristics, while correctly partitioning residual variance across levels and propagating uncertainty appropriately (Raudenbush & Bryk, 2002). The intraclass correlation coefficient (ICC), estimated from the unconditional HLM, quantifies the proportion of total outcome variance attributable to between-group differences, providing a direct empirical justification for the multilevel analytical strategy.

Despite the analytical power and theoretical superiority of HLM over conventional OLS and ANOVA approaches for nested data, the application of three-level hierarchical models—accommodating the three natural levels of students within classrooms within schools—remains relatively uncommon in the Indian educational research literature prior to 2010. Most existing Indian studies of school effectiveness and academic achievement employ either descriptive statistics or single-level regression models that treat classroom and school contexts as nuisance variables rather than as substantively interesting levels of the analysis. This methodological limitation obscures potentially important cross-level interactions—such as the differential effectiveness of experienced teachers depending on student-level ability composition—that can only be detected through explicit multilevel modelling.

The present study addresses these gaps by applying a rigorously specified three-level hierarchical linear model to a novel dataset of 1,478 secondary school students nested within 39 classrooms in 8 schools from a semi-urban district in southern India. The study pursues a systematic model-building strategy that progressively adds predictors at each hierarchical level, allowing the incremental contribution of each tier of the nesting structure to be quantified and statistically evaluated. A particular emphasis is placed on cross-level interactions—the moderation of student-level predictor effects by classroom- and school-level contextual variables—which represent the most theoretically distinctive capability of the HLM framework and have attracted increasing attention in the school effectiveness literature (Hox, 2002; Snijders & Bosker, 1999).

2. Review of Literature

2.1 Theoretical Foundations of Multilevel Modelling

The intellectual foundations of multilevel modelling can be traced to the general theory of mixed linear models developed by Henderson (1953) for applications in animal breeding and subsequently extended to a broad class of experimental designs by Hartley and Rao (1967). Their work on maximum likelihood estimation for mixed ANOVA models established the estimation framework—subsequently refined into restricted maximum likelihood (REML)—that underlies virtually all contemporary multilevel modelling software. Fisher's (1925) foundational treatment of experimental design and the analysis of variance established the conceptual framework within which the extensions to hierarchical structures were subsequently developed.

The behavioural science literature on hierarchical models was substantially advanced by Bock (1989), whose treatment of multivariate statistical methods for nested data structures highlighted the inadequacy of ignoring group-level dependencies in large-scale survey and experimental data. Bock's work drew explicit connections between the statistical theory of mixed models and the substantive research questions of behavioural science—particularly the decomposition of individual differences into

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

between-group and within-group components—that subsequently motivated the development of user-accessible multilevel modelling software.

Dempster, Rubin, and Tsutakawa (1981) provided a landmark methodological contribution by developing the EM algorithm-based approach to estimation in covariance component models, substantially improving the numerical stability and convergence of variance component estimation relative to earlier direct maximization approaches. Their work established the computational foundation for the reliable estimation of multilevel models in realistic research applications with unbalanced group sizes—a condition ubiquitous in educational and social science data.

Laird and Ware (1982) synthesized the growing literature on random effects models and provided a unified framework for their specification, estimation, and interpretation in the context of longitudinal data—a framework directly applicable to cross-sectional nested data through the reinterpretation of within-subject correlation as within-cluster correlation. Their treatment of the relationship between marginal and conditional models in the mixed effects framework clarified an important conceptual issue that has subsequently proven critical for the interpretation of multilevel model coefficients.

2.2 Hierarchical Linear Models in Educational Research

The application of hierarchical linear models to educational research was pioneered by Raudenbush and Bryk (1986), who demonstrated the inadequacy of both aggregated (school-level) and disaggregated (individual-level) analyses for studying school effects on student achievement, and established HLM as the appropriate analytical framework for linking student outcomes to school characteristics while correctly modelling the nested data structure. Their analysis of data from the High School and Beyond study provided compelling empirical evidence that between-school variance in student achievement is substantial and systematically related to school-level structural and process variables—a finding that ignited a major strand of educational effectiveness research.

Bryk and Raudenbush (1992) subsequently authored the definitive methodological treatment of hierarchical linear models for social and behavioural research, providing a comprehensive exposition of HLM specification, estimation, and interpretation across a range of research designs including two-level and three-level nested structures, longitudinal growth models, and cross-classified random effects models. This monograph established the intellectual framework and analytic conventions that have shaped educational HLM research for two decades.

Goldstein (1987) developed the multilevel modelling framework independently in the British educational statistics tradition, producing a parallel treatment that emphasized the connections between HLM and the general theory of linear models and that has been particularly influential in European educational research. Goldstein (2003) provided an updated and comprehensive treatment of multilevel statistical models covering a wide range of applications including normal response models, binary and ordered categorical responses, multivariate outcomes, and cross-classified structures.

Snijders and Bosker (1999) provided an accessible and rigorous treatment of multilevel analysis with extensive coverage of model specification, parameter interpretation, sample size planning, and model diagnostics, including influential discussions of how to determine the optimal allocation of sampling resources across hierarchical levels and how to assess the practical significance of ICC values. Their treatment of cross-level interactions as formal tests of contextual moderation effects has been particularly influential for school effectiveness research.

Hox (2002) extended the multilevel modeling literature with comprehensive coverage of three-level models, growth curve models, and models for non-continuous outcomes, and provided detailed guidance on the interpretation of random slope variance and covariance components that is directly applicable to the present study. His discussion of model-building strategies—beginning with the unconditional model and progressively adding predictors—provides the template followed in the present investigation.

2.3 School and Classroom Effects on Academic Achievement

Coleman et al. (1966), in their landmark 'Equality of Educational Opportunity' report, provided early large-scale evidence that family socioeconomic background accounts for the majority of variance in student achievement, with school characteristics contributing relatively modestly after controlling for student background. While methodologically limited by the use of aggregate-level analyses that ignored the nested structure of the data, the Coleman Report established the substantive agenda—the relative importance of school versus family factors—that has motivated much subsequent research.

Hanushek (1979) extended the school effectiveness literature with an influential econometric analysis of the production function for education, documenting that teacher quality—particularly teacher experience and academic credentials—is among the most important school-level determinants of student achievement gains. Hanushek's work established the human capital framework for studying school effectiveness that remains highly influential in educational economics research.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Creemers (1994) developed a comprehensive theory of educational effectiveness proposing that effective instruction at the classroom level—including curriculum quality, teacher behavior, and grouping procedures—mediates the relationship between school-level organizational factors and student-level learning outcomes. His multilevel theory of educational effectiveness provided the theoretical framework for numerous subsequent HLM-based studies of school and classroom effects.

Scheerens and Bosker (1997) provided a comprehensive meta-analytic review of the school effectiveness literature, synthesizing findings from hundreds of empirical studies and documenting that between-school variation in student achievement—as measured by the ICC—typically falls in the range of 10%–25% in developed-country educational systems, with somewhat higher values reported for developing-country settings. Their review highlighted the importance of controlling for student socioeconomic background when estimating school effects and provided evidence that instructional quality at the classroom level is a more proximal predictor of student outcomes than school-level structural variables.

Lee and Bryk (1989) conducted an influential HLM analysis of Catholic and public school effects on mathematics achievement in the United States, documenting that Catholic schools exhibit lower between-school variance in achievement—a finding attributed to their more academically focused and socially equitable organizational climate. Their analysis demonstrated the capacity of HLM to reveal cross-level interactions that are invisible in single-level analyses, establishing the paradigmatic template for subsequent school type comparative studies.

2.4 Indian Educational Context

Drèze and Sen (2002) provided a comprehensive assessment of Indian educational attainment, documenting persistent inequalities in school quality and student outcomes across socioeconomic, geographic, and caste-based population groups. Their analysis highlighted the structural impediments to educational effectiveness in the Indian context—including inadequate infrastructure, teacher absenteeism, and the urban–rural digital divide—that constitute important school-level moderators of the effectiveness of instructional investments. Their critique of aggregate educational statistics for masking within-school and between-school inequalities motivates the multilevel analytical approach adopted in the present study.

Fuller and Clarke (1994) reviewed the literature on school effects in developing countries, documenting that while the magnitude of school effects on student achievement is generally larger in developing than in developed countries—reflecting the greater heterogeneity in school quality—the specific school-level factors that predict effectiveness differ substantially across national contexts. Their review identified teacher characteristics, instructional materials availability, and school management type as among the most consistent cross-national predictors of school effectiveness in developing-country settings.

Govinda and Varghese (1993) provided an early systematic analysis of school quality variation in India, documenting substantial between-school heterogeneity in physical infrastructure, teacher qualifications, instructional practices, and student performance outcomes. Their analysis, while not employing multilevel modeling, provided the empirical backdrop for subsequent HLM-based studies of Indian school effectiveness and established the relevance of school management type (government versus private) as a potentially important school-level predictor of student achievement.

Kirk (1982) provided an authoritative treatment of experimental design procedures for the behavioral sciences that established the statistical framework—including the theory of nested, crossed, and mixed designs—within which hierarchical linear models for educational data are most naturally situated. His discussion of the properties of random effects models and the conditions under which between-group variance components are estimable provided foundational methodological guidance for educational HLM practitioners.

3. Research Gap

The foregoing review reveals four substantive gaps in the existing literature that the present study is designed to address.

First, although two-level HLM (students within schools) has been applied in several Indian educational studies, the extension to a three-level model (students within classrooms within schools) has received minimal attention in the Indian research literature prior to 2010. This omission is substantively important because classrooms—the immediate instructional environment—constitute the most proximal contextual level for student learning, and collapsing the classroom and school levels into a single random effect confounds classroom-specific instructional variation with school-level organizational and resource variation.

Second, the systematic examination of cross-level interactions in the Indian educational context—particularly the moderation of student-level effects by teacher characteristics and school-level SES composition—has not been systematically investigated using properly specified HLM frameworks. The detection of such interactions requires the full apparatus of the multilevel model and cannot be achieved through conventional regression or ANOVA approaches.

Third, the literature on Indian school effectiveness lacks a formal model comparison strategy based on likelihood ratio tests and information criteria that allows the incremental contribution of each hierarchical level to be quantified and statistically evaluated. Such a strategy is essential for determining whether the more complex three-level model provides a significantly better fit to the data than simpler specifications, thereby empirically justifying the additional model complexity.

Fourth, the treatment of unbalanced classroom and school sizes—a ubiquitous feature of real educational data in India—through fixed-effects ANOVA approaches introduces well-documented bias in variance component estimates. The present study employs REML estimation, which provides unbiased variance component estimates for unbalanced nested data, addressing this methodological limitation of prior work.

4. Research Objectives

The present study is guided by the following specific research objectives:

- To estimate the intraclass correlation coefficient (ICC) at both the classroom and school levels from an unconditional three-level HLM, thereby quantifying the proportion of total variance in student mathematics achievement attributable to classroom and school membership respectively.
- To identify statistically significant student-level predictors (prior academic performance, socioeconomic status, motivation, and gender) of mathematics achievement within the three-level HLM framework, with correctly estimated standard errors that account for the nested data structure.
- To assess the significance of classroom-level factors (teacher experience, teacher qualification, and class size) as predictors of classroom-mean mathematics achievement after controlling for student-level characteristics.
- To examine school-level structural and contextual variables (school management type, school mean SES, and school location) as predictors of between-school variation in mathematics achievement.
- To test formally for cross-level interactions between student-level predictors and classroom- or school-level moderating variables, identifying conditions under which the effects of student-level characteristics on achievement vary systematically across classroom and school contexts.

- To compare the fit of a sequence of progressively specified HLM models using likelihood ratio tests and information criteria, thereby evaluating the incremental explanatory contribution of each hierarchical level.

5. Hypotheses

Hypothesis Set 1: Variance Partitioning

H01: The intraclass correlation coefficient at the school level is zero; that is, school membership does not account for any systematic variance in student mathematics achievement.

Ha1: The intraclass correlation coefficient at the school level is significantly greater than zero, indicating meaningful between-school variation in mathematics achievement.

Hypothesis Set 2: Student-Level Predictors

H02: Prior academic performance, student SES, and student motivation do not significantly predict mathematics achievement within schools after accounting for school membership.

Ha2: Prior academic performance, student SES, and student motivation each significantly and positively predict individual mathematics achievement within the multilevel framework.

Hypothesis Set 3: Classroom-Level Predictors

H03: Teacher experience and class size do not significantly predict classroom-mean mathematics achievement after controlling for student composition.

Ha3: Teacher experience positively and class size negatively predicts classroom-mean mathematics achievement after controlling for student-level characteristics.

Hypothesis Set 4: School-Level Predictors

H04: School management type and school mean SES do not significantly explain between-school variation in mathematics achievement.

Ha4: Private school management and higher school mean SES are associated with significantly higher mean mathematics achievement at the school level.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

Hypothesis Set 5: Cross-Level Interactions

H05: The effect of student-level prior performance on mathematics achievement does not differ significantly across classrooms with different levels of teacher experience.

Ha5: Teacher experience significantly moderates the within-school effect of prior academic performance, such that students with lower prior performance benefit more from experienced teachers.

Hypothesis Set 6: Model Fit

H06: The three-level HLM does not provide significantly better fit to the data than a single-level OLS regression or a two-level model.

Ha6: The three-level HLM with cross-level interactions provides significantly better fit than single-level and two-level nested specifications, as evidenced by likelihood ratio tests and information criteria.

6. Research Methodology

6.1 Research Design

The present study employs an observational, cross-sectional research design with a quantitative analytical framework grounded in the theory of mixed linear models and hierarchical linear modeling. The design is analytical rather than experimental in the conventional sense: students, classrooms, and schools constitute naturally occurring nested groups, and no random assignment of students to treatment conditions was performed. The nesting structure of the data—students within classrooms within schools—constitutes a three-level nested design in the sense of Kirk (1982), with classrooms constituting the primary nested units at Level 2 and schools constituting the higher-order grouping units at Level 3.

6.2 Population and Sampling

The target population comprises secondary school students (Grade 9 and 10) enrolled in government and private schools in a semi-urban district in southern India. A cluster sampling design was employed, with schools as the primary sampling unit. Eight schools were selected to represent the geographic and school management type diversity of the district: three urban government, two urban private, two peri-urban government, and one rural private school. Within each selected school, all available Grade 9 and 10 classrooms (4–6 per school) were included, yielding 39 classrooms in total. All

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

students enrolled in the selected classrooms who were present on the day of data collection and who provided complete responses to the survey instruments were included, resulting in a final analyzed sample of 1,478 students.

6.3 Instruments and Measures

Mathematics achievement was assessed using a standardized 50-item diagnostic test aligned with the state secondary school mathematics curriculum, covering algebra, geometry, statistics, and arithmetic reasoning. The test was administered under standardized conditions by trained field investigators. Prior academic performance was measured using the student's score from the most recent school-term examination, available from school records. Student socioeconomic status was operationalized using a composite index derived from parental education, parental occupation, and household asset ownership, standardized to mean zero and unit standard deviation across the full sample. Student motivation was measured using a five-item Likert scale adapted from validated instruments in the school motivation literature, with responses ranging from 1 (strongly disagree) to 5 (strongly agree), scoring yielding a mean scale score. Teacher experience was measured in years of full-time teaching at the secondary level. Class size was the total enrollment in the class section as of the census date.

6.4 Statistical Model Specification

Following Raudenbush and Bryk (2002) and Hox (2002), a three-level hierarchical linear model was specified. At Level 1 (student level), mathematics achievement (Y) for student i in classroom j in school k is modeled as:

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk}(PRIOR)_{ijk} + \beta_{2jk}(SES)_{ijk} + \beta_{3jk}(MOTIV)_{ijk} + r_{ijk} \quad \ldots (6.1)$$

where $r_{ijk} \sim N(0, \sigma^2)$ is the Level-1 residual.

At Level 2 (classroom level), each intercept and slope from equation (6.1) becomes a dependent variable predicted by classroom-level characteristics:

$$\beta_{0jk} = \gamma_{00k} + \gamma_{01}(CLASSSIZE)_{jk} + \gamma_{02}(TCHEXP)_{jk} + u_{0jk} \quad \ldots (6.2)$$

$$\beta_{1jk} = \gamma_{10k} + \gamma_{11}(TCHEXP)_{jk} + u_{1jk} \quad \ldots (6.3)$$

$$\beta_{2jk} = \gamma_{20k} + u_{2jk} \quad \ldots (6.4)$$

$$\beta_{3jk} = \gamma_{30k} \quad \ldots (6.5)$$

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362
IJARST

where u_{0jk}, u_{1jk}, u_{2jk} are Level-2 random effects assumed multivariate normally distributed with covariance matrix T_2.

At Level 3 (school level), the classroom-level intercepts from equation (6.2) are predicted by school-level variables:

$\gamma_{00k} = \delta_{00} + \delta_{01}(SCHL\_SES)_k + \delta_{02}(SCHL\_TYPE)_k + v_{0k}$ … (6.6)

$\gamma_{20k} = \delta_{20} + \delta_{21}(SCHL\_SES)_k + v_{2k}$ … (6.7)

where v_{0k} and v_{2k} are Level-3 random effects assumed bivariate normally distributed with covariance matrix T_3.

6.5 Estimation and Model Comparison

All models were estimated using Restricted Maximum Likelihood (REML), which yields unbiased variance component estimates for unbalanced nested data (Harville, 1977). Fixed effects were estimated using generalized least squares with REML-estimated covariance matrices. Model comparison was conducted using likelihood ratio tests for nested model comparisons (with full maximum likelihood to enable cross-model comparison on the same likelihood scale) and AIC and BIC information criteria for non-nested comparisons. The ICC was computed from the null model as $\rho = \tau_{00} / (\tau_{00} + \sigma^2)$, where τ_{00} is the between-group (school-level) variance and σ^2 is the within-group (student-level) residual variance.

6.6 Ethical Considerations and Limitations

Ethical approval was obtained from the relevant institutional review board. Written informed consent was obtained from school principals, teachers, and (for student data) from parents or legal guardians of participating students. Individual student data are reported only in aggregate form. The primary limitations of the present study include: the cross-sectional design precludes causal inference about the effects of school and classroom characteristics on student achievement; the sample of eight schools, while sufficient for estimating school-level variance components, provides limited statistical power for estimating school-level fixed effects; and the reliance on a single district limits the generalizability of the findings to other regional and linguistic contexts within India.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in
IJARST

7. Data Analysis and Interpretation

7.1 Sample Characteristics

Table 1 presents the structural characteristics of the eight schools and their constituent classrooms. School sizes range from 127 to 241 enrolled students in the target grades, with class sizes averaging 38.4 students. The overall sample is balanced with respect to gender (50.4% female). Government schools account for 62.5% of sampled schools and approximately 54.5% of sampled students, consistent with the enrollment distribution in the target district.

Table 1: Sample Design Summary — Schools, Classrooms, and Student Distribution

School (Level-2 Unit)	Location	School Type	No. of Classes	Class Size (avg)	Total Students	Female %
School 1	Urban	Govt.	6	38.2	229	51.5
School 2	Urban	Private	5	34.6	173	48.9
School 3	Peri-urban	Govt.	6	40.1	241	52.3
School 4	Rural	Govt.	4	42.5	170	49.4
School 5	Rural	Private	4	31.8	127	53.1
School 6	Urban	Private	5	33.4	167	47.3
School 7	Peri-urban	Govt.	5	39.7	199	50.8
School 8	Rural	Govt.	4	43.1	172	48.2
Total / Mean	—	—	39	38.4	1,478	50.4

Note: Peri-urban schools are classified as non-urban by district geographic designation but are located within 15 km of a major urban center.

7.2 Descriptive Statistics

Table 2 reports descriptive statistics for all variables included in the HLM analyses. Mathematics achievement scores exhibit a near-normal distribution (skewness = −0.18; kurtosis = 2.84) with a mean of 58.43 and a standard deviation of 14.27, indicating adequate distributional spread for the proposed analyses. The student SES index, standardized across the full sample, shows a slight positive skew (0.12) reflecting the modest upper-SES overrepresentation in the private school subsample. Teacher experience at the classroom level averages 9.63 years and varies substantially across classrooms (range: 1–22 years), providing adequate variation to detect classroom-level teacher effects. School mean SES ranges from −0.94 to +0.87 standard deviation units across the eight schools, reflecting meaningful socioeconomic stratification in the sample.

Table 2: Descriptive Statistics for All Study Variables

Variable	N	Mean	SD	Min	Max	Skew	Kurt
Math Achievement Score	1478	58.43	14.27	22	98	−0.18	2.84
Prior Academic Performance	1478	54.61	13.89	18	95	−0.09	2.76
Socioeconomic Status Index	1478	0.00	1.00	−2.41	2.83	0.12	2.93
Student Motivation Scale	1478	3.74	0.81	1.0	5.0	−0.31	3.12
Teacher Experience (yrs)	39	9.63	4.71	1	22	0.41	2.67
Class Size	39	38.41	4.53	28	47	−0.07	2.81
School Mean SES	8	0.00	0.63	−0.94	0.87	−0.04	1.88

Note: Student-level variables (rows 1–4) are reported for N = 1,478 students; classroom-level variables (rows 5–6) are reported for N = 39 classrooms; school-level variables (row 7) are reported for N = 8 schools. SES = socioeconomic status index (standardized).

7.3 Variance Partitioning — Unconditional Model

Table 3 presents the results of the null (unconditional) model and the sequence of progressively specified models. The null model decomposes total variance in mathematics achievement into student-level ($\sigma^2 = 148.73$), and school-level ($\tau_{00} = 51.34$) components, yielding an ICC of 0.257. This indicates that approximately 25.7% of total variance in student mathematics achievement is attributable to systematic between-school differences—a value substantially above the threshold of 10% commonly cited as justifying multilevel analysis (Snijders & Bosker, 1999) and consistent with the higher ICCs typically reported in developing-country educational settings (Scheerens & Bosker, 1997). H01 is unequivocally rejected.

Progressive model building substantially reduces the Level-2 (school-level) residual variance from 51.34 in the null model to 14.28 in the full model—a reduction of 72.2%—while reducing Level-1 residual variance more modestly (from 148.73 to 117.61, a reduction of 21.0%). This pattern reflects the disproportionate capacity of school-level predictors to explain between-school achievement differences and is consistent with the compositional and contextual mechanisms emphasized in the school effectiveness literature (Lee & Bryk, 1989; Creemers, 1994).

Table 3: Variance Decomposition and ICC Across Progressive HLM Specifications

Model	Level-1 Variance (σ^2)	Level-2 Variance (τ_{00})	ICC (ρ)	−2LL
Null (Unconditional) Model	148.73	51.34	0.257***	11042.6
Model 1: Student Predictors Only	121.46	44.81	0.270	10731.4
Model 2: + Class-Level Predictors	119.87	31.22	0.207	10612.8
Model 3: + School-Level Predictors	118.93	18.47	0.134	10489.3
Model 4: Full Cross-Level Interaction	117.61	14.28	0.108	10421.7

*Note: ICC = intraclass correlation = $\tau_{00} / (\tau_{00} + \sigma^2)$. *** $p < .001$. All models estimated by REML; −2LL values from full ML estimation for comparability.*

7.4 Fixed Effects — Full HLM

Table 4 presents the fixed effects estimates from the fully specified Model 4, which includes student-level predictors, classroom-level predictors, school-level predictors, and cross-level interactions. All parameter estimates are based on REML estimation with empirical Bayes standard errors.

At the student level, prior academic performance ($\gamma_{10} = 0.531$, $p < .001$), student SES ($\gamma_{20} = 3.29$, $p < .001$), and student motivation ($\gamma_{30} = 2.87$, $p < .001$) are all significant positive predictors of mathematics achievement, consistent with the extant literature on student-level determinants of academic success (Hanushek, 1979; Scheerens & Bosker, 1997). These findings support rejection of H02.

At the classroom level, class size exerts a significant negative effect on mathematics achievement ($\gamma_{40} = -0.31$, $p = .008$), consistent with the hypothesis that larger classes impede individualized instruction and feedback (Coleman et al., 1966; Creemers, 1994). Teacher experience is a significant positive predictor of classroom mean achievement ($\gamma_{50} = 0.47$, $p = .018$), supporting Ha3.

At the school level, school mean SES ($\gamma_{01} = 4.17$, $p = .024$) and private school management ($\gamma_{02} = 3.82$, $p = .067$, marginally significant) are both associated with higher grand-mean mathematics achievement, supporting Ha4 and consistent with the school effectiveness findings reviewed by Scheerens and Bosker (1997) and the Indian-context evidence of Fuller and Clarke (1994).

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

Table 4: Fixed Effects Estimates — Full Three-Level HLM (Model 4)

Fixed Effect	Coeff. (γ)	SE	t-ratio	df	p-value	95% CI
For Intercept β_{0j}						
Grand Mean (γ_{00})	58.43***	1.84	31.75	7	<.001	[54.62, 62.24]
School Mean SES (γ_{01})	4.17**	1.39	3.00	6	0.024	[1.28, 7.06]
School Type: Private (γ_{02})	3.82*	1.71	2.23	6	0.067	[0.28, 7.36]
For Prior Perf. Slope β_{1j}						
Slope Intercept (γ_{10})	0.531***	0.041	12.95	37	<.001	[0.448, 0.614]
Teacher Exp. × Prior Perf. (γ_{11})	0.018*	0.008	2.25	37	0.030	[0.002, 0.034]
For SES Slope β_{2j}						
Slope Intercept (γ_{20})	3.29***	0.54	6.09	37	<.001	[2.19, 4.39]
School Mean SES × SES (γ_{21})	1.12*	0.47	2.38	6	0.054	[−0.06, 2.30]
For Motivation Slope β_{3j}						
Slope Intercept (γ_{30})	2.87***	0.48	5.98	37	<.001	[1.90, 3.84]
Class Size (γ_{40})	−0.31**	0.11	−2.82	37	0.008	[−0.53, −0.09]
Teacher Experience (γ_{50})	0.47*	0.19	2.47	37	0.018	[0.08, 0.86]

*Note: *** $p < .001$; ** $p < .01$; * $p < .05$. SE = empirical Bayes standard error. df for Level-1 effects = 37 (classrooms − parameters); df for school-level effects = degrees of freedom based on number of Level-3 units. γ notation follows Raudenbush and Bryk (2002).*

7.5 Cross-Level Interactions

Two significant cross-level interactions are identified in Model 4. First, teacher experience significantly moderates the within-school effect of prior academic performance on achievement (γ_{11} = 0.018, p = .030): the positive effect of prior performance is amplified in classrooms with more experienced teachers. Substantively, this indicates that experienced teachers are particularly effective at translating students' prior knowledge base into current achievement gains—consistent with Hanushek's (1979) view of teacher quality as a productivity-enhancing factor in the educational production function. This finding supports Ha5 and rejects H05.

Second, the effect of individual SES on achievement is moderated by school mean SES (γ_{21} = 1.12, p = .054, marginally significant): students with higher individual SES benefit somewhat more from

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

attending schools with a higher mean SES composition. This compositional effect is consistent with peer effects mechanisms described in the school effectiveness literature (Coleman et al., 1966; Lee & Bryk, 1989) and suggests that the equalization of school SES composition could partially compensate for individual-level socioeconomic disadvantage.

7.6 Random Effects

Table 5 presents the random effects estimates from Model 4. The significant variance in classroom intercepts (τ_{00} = 14.28, χ^2 = 47.31, $p < .001$) confirms that, after accounting for all modeled predictors, substantial systematic between-school variation in achievement remains unexplained by the included school-level variables. The significant variance in the prior performance slope (τ_{11} = 0.0031, χ^2 = 14.22, $p = .047$) and SES slope (τ_{22} = 0.841, χ^2 = 18.64, $p = .009$) indicates that the effects of prior performance and SES on student achievement vary significantly across schools—a substantive finding suggesting that the school context moderates the translation of individual background characteristics into academic outcomes in ways not fully captured by the modeled school-level predictors.

Table 5: Random Effects Estimates — Full Three-Level HLM (Model 4)

Random Effect	Variance Comp.	SD	χ^2	df	p-value
Intercept variance (τ_{00})	14.28	3.779	47.31	7	<.001
Prior Perf. slope var. (τ_{11})	0.0031	0.056	14.22	7	0.047
SES slope variance (τ_{22})	0.841	0.917	18.64	7	0.009
Intercept–Prior Perf. cov. (τ_{01})	−0.142	—	—	—	—
Level-1 residual (σ^2)	117.61	10.845	—	—	—

Note: Variance components estimated by REML. χ^2 statistics test the hypothesis that each variance component equals zero. τ_{01} = covariance between intercept and prior performance slope (not tested for significance separately).

7.7 Model Fit Comparison

Table 6 presents the model fit statistics for the sequence of HLM specifications. Each model increment provides a significantly better fit than its predecessor, as evidenced by the chi-square difference tests based on full ML estimation (all $\Delta\chi^2$ $p < .001$). The AIC and BIC also decline monotonically from

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

the null model through Model 4, confirming that the additional parameters introduced at each level are justified by the improvement in model fit. The full Model 4 is selected as the preferred specification on the basis of both statistical significance of the model comparisons and the substantive theoretical justification for all included predictors and interactions. These results provide strong empirical support for the rejection of H06.

Table 6: Model Fit Statistics — Sequential HLM Specifications

Model	Parameters	−2LL	AIC	BIC	Δχ² (df)
Null Model (M0)	3	11042.6	11048.6	11065.4	—
M1: Student-Level	9	10731.4	10749.4	10799.5	311.2*** (6)
M2: + Class-Level	13	10612.8	10638.8	10711.2	118.6*** (4)
M3: + School-Level	16	10489.3	10521.3	10609.0	123.5*** (3)
M4: + Cross-Level Interactions	19	10421.7	10459.7	10563.8	67.6*** (3)

*Note: Parameters = total number of estimated parameters (fixed effects + variance components). AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion. Δχ² = log-likelihood ratio test statistic relative to the preceding model; *** p < .001.*

8. Results and Discussion

The empirical results of the present study converge on several findings of both methodological and substantive significance. The central methodological finding is the confirmation that data from students nested within classrooms within schools in the Indian educational context exhibit substantial intraclass correlation (ICC = 0.257 at the school level), fully justifying the three-level HLM framework and establishing that conventional single-level OLS regression or standard one-way ANOVA would have produced severely underestimated standard errors and spuriously significant predictor effects. This finding is consistent with the general prediction of Bryk and Raudenbush (1992) and the meta-analytic evidence of Scheerens and Bosker (1997) that developing-country educational settings exhibit higher between-school variance than developed-country counterparts.

The identification of prior academic performance as the single most powerful student-level predictor (γ_{10} = 0.531) is consistent with the extensive literature documenting the predictive validity of prior achievement for subsequent learning outcomes (Hanushek, 1979; Raudenbush & Bryk, 2002). The

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

substantial positive effects of student SES (γ_{20} = 3.29) and motivation (γ_{30} = 2.87) confirm the well-established finding that socioeconomic background and motivational dispositions are among the most consistent and policy-relevant individual-level predictors of academic achievement, and that these effects persist after properly accounting for school and classroom clustering.

At the classroom level, the negative effect of class size (γ_{40} = −0.31 per additional student) and positive effect of teacher experience (γ_{50} = 0.47 per year of teaching) are consistent with the educational economics literature (Hanushek, 1979; Creemers, 1994) and provide direct empirical support for policy interventions targeting class size reduction and teacher retention in the Indian secondary school context. The relatively modest magnitude of the class size effect, however, suggests that class size reductions below a threshold of approximately 30–35 students may be necessary to produce educationally meaningful achievement gains—a threshold not approached in the current sample, where mean class size is 38.4.

The school-level finding that private school management is associated with approximately 3.82 additional score points in mathematics achievement (after controlling for school mean SES and student background) is consistent with the Indian-context evidence reviewed by Govinda and Varghese (1993) and Fuller and Clarke (1994) but is more modest than the school type effects documented by Lee and Bryk (1989) in the U.S. Catholic school context. The moderation of this effect by school mean SES underscores the compositional complexity of private school advantages in the Indian context, where the private school sector enrolls disproportionately high-SES students whose academic performance is enhanced by both school quality and peer composition effects.

The cross-level interaction between teacher experience and prior academic performance is the most theoretically novel finding of the present study. Its positive sign (γ_{11} = 0.018) indicates that the academic productivity of prior knowledge—the rate at which prior performance translates into current achievement—is amplified in classrooms with more experienced teachers. This is consistent with a teaching quality interpretation in which experienced teachers possess superior pedagogical content knowledge that enables them to build more effectively on the academic foundations students bring to the classroom. From an equity perspective, this interaction implies that less experienced teachers—disproportionately deployed in low-performing schools in the Indian system—may be less effective at leveraging the academic potential of their students, contributing to a self-reinforcing cycle of educational disadvantage.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

9. Conclusion

The present study has developed and estimated a three-level hierarchical linear model for student mathematics achievement using data from 1,478 students nested within 39 classrooms in 8 schools in southern India. The unconditional null model confirmed that approximately 25.7% of total variance in mathematics achievement is attributable to between-school differences—a magnitude that fully justifies the multilevel analytical framework and establishes the inadequacy of single-level regression or ANOVA approaches for this class of educational data.

The full three-level HLM identified significant predictors at each hierarchical level: prior academic performance, SES, and motivation at the student level; class size and teacher experience at the classroom level; and school mean SES and management type at the school level. The model explained 72.2% of between-school variance relative to the null model. Two significant cross-level interactions—between teacher experience and student prior performance, and between school mean SES and individual SES—reveal that the effects of student-level characteristics on achievement are systematically moderated by the classroom and school context in theoretically meaningful ways.

These findings collectively demonstrate the analytical superiority of the three-level HLM over conventional ANOVA and OLS approaches for nested educational data, contribute to the empirical literature on school effectiveness in the developing-country context, and provide actionable evidence for educational policy decisions regarding teacher deployment, class size management, and school resource allocation in India.

Future research should extend the present framework to longitudinal (growth model) specifications to capture the dynamic evolution of school and teacher effects over academic years; to larger, nationally representative school samples to enhance the statistical power of school-level fixed effect estimates; and to the examination of additional classroom-level instructional process variables—such as instructional time allocation, homework frequency, and formative assessment practices—that are theoretically central to the Creemers (1994) educational effectiveness model but were not available in the present dataset.

10. References

Bock, R. D. (1989). Multilevel analysis of educational data. Academic Press.

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

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). Equality of educational opportunity. U.S. Government Printing Office.

Creemers, B. P. M. (1994). The effective classroom. Cassell.

Dempster, A. P., Rubin, D. B., & Tsutakawa, R. K. (1981). Estimation in covariance component models. Journal of the American Statistical Association, 76(374), 341–353.

Drèze, J., & Sen, A. (2002). India: Development and participation. Oxford University Press.

Fisher, R. A. (1925). Statistical methods for research workers. Oliver and Boyd.

Fuller, B., & Clarke, P. (1994). Raising school effects while ignoring culture? Local conditions and the influence of classroom tools, rules, and pedagogy. Review of Educational Research, 64(1), 119–157.

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

Goldstein, H. (2003). Multilevel statistical models (3rd ed.). Arnold.

Govinda, R., & Varghese, N. V. (1993). Quality of primary schooling in India: A case study of Madhya Pradesh. International Institute of Educational Planning / NIEPA.

Hanushek, E. A. (1979). Conceptual and empirical issues in the estimation of educational production functions. Journal of Human Resources, 14(3), 351–388.

Hartley, H. O., & Rao, J. N. K. (1967). Maximum likelihood estimation for the mixed analysis of variance model. Biometrika, 54(1–2), 93–108.

Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. Journal of the American Statistical Association, 72(358), 320–338.

Henderson, C. R. (1953). Estimation of variance and covariance components. Biometrics, 9(2), 226–252.

Hox, J. J. (2002). Multilevel analysis: Techniques and applications. Lawrence Erlbaum Associates.

Kirk, R. E. (1982). Experimental design: Procedures for the behavioral sciences (2nd ed.). Brooks/Cole.

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. Biometrics, 38(4), 963–974.

Lee, V. E., & Bryk, A. S. (1989). A multilevel model of the social distribution of high school achievement. Sociology of Education, 62(3), 172–192.

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

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

Scheerens, J., & Bosker, R. J. (1997). The foundations of educational effectiveness. Pergamon Press.

Snijders, T. A. B., & Bosker, R. J. (1999). Multilevel analysis: An introduction to basic and advanced multilevel modeling. Sage.

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

Longford, N. T. (1987). A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. Biometrika, 74(4), 817–827.

Prosser, R., Rasbash, J., & Goldstein, H. (1991). ML3 software for three-level analysis: Users' guide for V.2. Institute of Education, University of London.

Willms, J. D. (1992). Monitoring school performance: A guide for educators. Falmer Press.

Lockheed, M. E., & Komenan, A. (1989). Teaching quality and student achievement in Africa: The case of Nigeria and Swaziland. Teaching and Teacher Education, 5(2), 93–113.

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