DOCUMENT RESUME

ED 380 475 TM 022 777

AUTHOR	Arnold, Carolyn L.
TITLE	Using HLM and NAEP Data To Explore School Correlates of 1990 Mathematics and Geometry Achievement in Grades 4, 8, and 12: Methodology and Results. Research and Development Report.
INSTITUTION	MPR Associates, Berkeley, CA.
SPONS AGENCY	National Center for Education Statistics (ED), Washington, DC.
REPORT NO	ISBN-0-16-045449-2; NCES-95-697
PUB DATE	Jan 95
NOTE	213p.
AVAILABLE FROM	U.S. Government Printing Office, Superintendent of Documents, Mail Stop: SSOP, Washington, DC 20402-9328.
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Statistical Data (110)
EDRS PRICE	MF01/PC09 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Correlation; Elementary Secondary Education; Ethnicity; *Geometry; Grade 4; Grade 8; Grade 12; Institutional Characteristics; *Mathematics Achievement; Outcomes of Education; Prediction; Racial Differences; *Research Methodology; Sex Differences; Tables (Data)
IDENTIFIERS	*Hierarchical Linear Modeling; *National Assessment of Educational Progress

ABSTRACT

This report illustrates the use of hierarchical linear models (HLM) with National Assessment of Educational Progress (NAEP) data to identify school and other correlates of student achievement. Based on an analysis of the 1990 NAEP mathematics achievement data for 4th, 8th, and 12th graders in public schools, this study is part of an ongoing exploratory effort to demonstrate the potential usefulness of HLM with NAEP data. The focus of the report is on the methodology of using HLM with NAEP data, and results of the study are presented as an illustration of the methodology. HLM accurately models the multilevel nature of the data and enables student-level outcomes such as gender and race/ethnicity to be predicted as a function of school-level factors. Several types of HLM analysis were conducted on 1990 NAEP data to predict achievement outcomes in mathematics and geometry, predicting average achievement between schools, the gender gap, and the race/ethnicity gap. The HLM methods worked well to explain variations in achievement but less well for the gender gap and race/ethnicity gaps. Results are discussed in the context of improving the usefulness of NAEP data. Data are presented in 66 tables, with 42 tables of supporting data in an appendix. (SLD)

NATIONAL CENTER FOR EDUCATION STATISTICS

Research and Development Report January 1995

Using HLM and NAEP Data to Explore School Correlates of 1990 Mathematics and Geometry Achievement in Grades 4, 8, and 12: Methodology and Results

U.S. Department of Education
Office of Educational Research and Improvement NCES 95-697

NATIONAL CENTER FOR EDUCATION STATISTICS

Research and Development Report January 1995

Using HLM and NAEP Data to Explore School Correlates of 1990 Mathematics and Geometry Achievement in Grades 4, 8, and 12: Methodology and Results

Carolyn L. Arnold
MPR Associates, Inc.

U.S. Department of Education
Office of Educational Research and Improvement NCES 95-697

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
Sharon P. Robinson
Assistant Secretary

National Center for Education Statistics
Emerson J. Elliott
Commissioner

National Center for Education Statistics

"The purpose of the Center shall be to collect, analyze, and disseminate statistics and other data related to education in the United States and in other nations."—Section 406(b) of the General Education Provisions Act, as amended (20 U.S.C. 1221e–1).

January 1995

Contact:
Alex Sedlacek
(202) 219–1734

For sale by the U.S. Government Printing Office
Superintendent of Documents, Mail Stop: SSOP, Washington, DC 20402-9328
ISBN 0-16-045443-2

Foreword

The Research and Development (R&D) series of reports has been initiated

1) To share studies and research that are developmental in nature. The results of such studies may be revised as the work continues and additional data become available.

2) To share results that are, to some extent, on the "cutting edge" of methodological developments. Emerging analytical approaches and new computer software developments often permit new, and sometimes controversial analysis to be done. By participating in "frontier research," we hope to contribute to the resolution of issues and improved analysis.

3) To participate in discussions of emerging issues of interest to educational researchers, statisticians, and the Federal statistical community in general. Such reports may document workshops and symposiums sponsored by NCES that address methodological and analytical issues or may share and discuss issues regarding NCES practice, procedures, and standards.

The common theme in all these goals is that these reports present results or discussion that do not reach definitive conclusions at this point in time, either because the data are tentative, the methodology is new and developing, or the topic is one on which there are divergent views. Therefore the techniques and inferences made from the data are tentative and are subject to revision. To facilitate the process of closure on the issues, we invite comment, criticism, and alternatives to what we have done. Such responses should be addressed to:

Emerson Elliott
Commissioner
National Center for Education Statistics
555 New Jersey Ave. NW
Washington, D.C. 20208

Acknowledgments

This study was supported by a grant from the NAEP Data Reporting Program, Education Assessment Division, National Center for Education Statistics, U.S. Department of Education. That grant allowed the author to extend her knowledge of HLM and NAEP and apply it to the latest NAEP data. Throughout the period of the grant, the grant administrator, Deborah Sedlacek, facilitated the project in many tangible and intangible ways. At MPR Associates, Phillip Kaufman served as a supportive and wise colleague. The study could not have been completed without the brilliant and careful research assistance of Aziza Khazzoom. Her assistance covered everything from early theorizing to tedious data cleaning to investigating puzzling software bugs. Her attention to the detail of all the variables, programs, results, and software were appreciated far more than was expressed at the time, and enabled the study to proceed accurately. Stephen Raudenbush of Michigan State University was available for thorny HLM questions, and Frank Jenkins of ETS provided valuable NAEP information and statistical insights. Mark Wilson of U.C. Berkeley provided local statistical and research advice. In the software project, Charles Breaux produced complicated code under exacting conditions and remained congenial at all times. Richard Congden of Brandeis University was very generous with his time and programming as we set about modifying his HLM code. Cathy Gilbert helped at crucial junctures by supervising the table formatting and attending to other logistical issues with the help of numerous support staff. Finally, the author would like to express appreciation to Gary Hoachlander and Susan Choy of MPR Associates for enabling the author to make a smooth transition from MPR employee to MPR consultant during the last months of this grant.

Executive Summary

Overview of report

This report illustrates the use of hierarchical linear models (HLM) with NAEP data to identify school and other correlates of student achievement. Based on an analysis of the 1990 NAEP mathematics achievement data for 4th, 8th, and 12th graders in public schools, this study is part of an ongoing exploratory effort to demonstrate the potential usefulness of HLM, a state-of-the-art statistical procedure with NAEP, a complex data set.

The focus of the report is on the methodology of using HLM with NAEP data, and the results of the study are presented as an illustration of that methodology. Due to the exploratory nature of the study and limitations of the data, policy changes are not recommended on the basis of this report. Instead, researchers are encouraged to use this analysis as the basis for an understanding of the procedures and questions involved in using NAEP data and hierarchical linear models for school effectiveness studies.

Overview of HLM models

HLM allows the examination of associations among multi-level, nested data such as students within schools by estimating simultaneous linear equations at the student level within schools and the school level between schools. HLM models explain student and school variation in achievement scores, using both student- and school-level variables as explanatory variables, while accounting for the variance at each level. Thus, HLM accurately models the multi-level nature of the data. In addition, HLM enables student-level outcomes such as gender and race-ethnicity differences in achievement to be predicted as a function of school-level factors.

In this study, several types of HLM analyses were conducted on the 1990 NAEP data to predict achievement outcomes in overall mathematics and in geometry, one of the higher-level subscales of mathematics achievement. The student characteristics used to predict mathematics and geometry achievement within schools were gender, race-ethnicity, socioeconomic status (SES), and coursetaking (in grades 8 and 12). Six groups of school characteristics were used to predict school-level outcomes: student body characteristics; school fiscal, physical, and staff resources; classroom instructional methods; and three components of campus climate (attitudes toward mathematics, student safety and behavior, and academic expectations). Within each grade, and separately for mathematics and geometry, HLM models estimated the association of the six groups of school characteristics with 1) the average achievement within schools; 2) the association of gender with achievement within schools, or the achievement gap between female students and male students within schools; and 3) the association of race-ethnicity with achievement within schools, or the achievement gap between African-American, Hispanic, and Native American students and white and Asian-American students within schools. In addition, two other exploratory analyses were conducted that examined alternative models for predicting the gender and race-ethnicity achievement gaps. The literature on school effects and on differences in mathematics achievement by gender and race-ethnicity provided the theoretical framework for these models.

Considerations in developing HLM models with NAEP data

In building the 1990 HLM models, five major considerations were taken into account. In addition, specific NAEP variables were selected for the models based on the factors of variability, interpretive meaning, collinearity, and missing values. The major considerations in building the models were:

1) Duplicating an earlier study. This study builds upon and improves a multi-level analysis of school effects on math and science achievement was conducted on the 1986 NAEP data (see C. Arnold, P. Kaufman, and D. Sedlacek 1992 in chapter 1). That analysis was the first time HLM had been applied to the multi-level NAEP data, and it illustrated the value of using HLM when analyzing NAEP. In this study, two of the six 1990 models are similar to 1986 models, and the 1990 results from these models are compared to the 1986 results.

2) Developing new models. New models were developed that focused more on the classroom instruction and school climate factors that are believed to be related to student achievement and to gender and race-ethnicity differences in achievement.

3) Maximizing the use of conditioning variables. Conditioning variables are those student, teacher, and school-level variables that were used to estimate the student plausible values of the achievement scores. The 1990 study was designed to determine whether using more conditioning variables would make a difference in the results and whether the conditioning variables would perform better than the nonconditioning variables. The 1990 NAEP used many more conditioning variables than the 1986 NAEP, and in developing the models for this analysis, conditioning variables were used wherever possible.

4) Choosing between random and fixed slope models. This consideration affected which student-level variables would be modeled and which would be used as control variables. For the major analyses of math and geometry, all three grades, and all six school-level models, a standard student-level model was developed. However, the alternative models explored using different combinations of fixed and random slopes.

5) Deciding whether to estimate separate or combined models. This study developed and tested six separate between-school models, rather than testing one general model of all the variables or testing six models and combining the significant variables into one final model. Student body characteristics were included as control variables in each model. However, combined models can also be valuable, and a recommended second step might be to choose theoretically important (not necessarily significant) variables from each model to combine into a general model.

Results: predicting average achievement between schools

The final models identified several school characteristics that significantly predicted average achievement and the gender and race-ethnicity gaps within schools. Most of the significant results were based on variables that were both new to the models and were conditioning variables. The results include the following:

Average achievement in the schools varied widely between the schools, and the HLM models were designed to model and explain this variation. In this study, student body characteristics were associated with average achievement, as they were in the 1986 study. In all three grades, schools with a higher percentage of African-American students in the school averaged lower math and geometry achievement, while schools with higher SES levels averaged higher achievement. Unlike the 1986 study, school resources were generally not associated with achievement. However, in grade 4, schools with higher percentages of students using computers in the school averaged higher math achievement.

Several classroom instructional methods used in math classes in the schools were associated with average achievement. Consistently, more time spent in doing problems from textbooks was associated with higher math and geometry achievement in grades 4 and 8,. In addition, working with objects (rulers, blocks, shapes, and solids) was positively associated with geometry achievement in grade 4. Using calculators in math classes in grade 4 was negatively associated with math achievement, while using computers in math classes in grade 8 was negatively associated with math achievement and geometry achievement. However, both using calculators and writing math proofs in math classes were associated with higher math and geometry achievement in grade 12.

School climate measures explained some of the variation in achievement. In one model, schools in which a larger proportion of 4th or 8th grade students held positive views about females and math averaged higher math achievement for all students in these grades. However, this was only significant when controlling for academic expectations of a school. More consistently, the more that 8th and 12th grade students felt they were good at and liked math, the higher was their average achievement in mathematics in grade 8 and in geometry in grade 12. Higher academic expectations were also associated with higher achievement in grades 8 and 12. In addition, in grade 12, schools with higher levels of disruption in the classroom averaged lower math and geometry achievement.

Results: predicting the gender gap

Within schools in grades 4 and 8, little or no average gender differences were found in math or geometry achievement, controlling for race-ethnicity, SES, and coursetaking. However, a wide range of gender differences existed between schools in these grades, and models were developed to explain that variation. In contrast, in grade 12, females averaged 2.5 points achievement points lower than males in math and 6.5 points lower than males in geometry, and these differences varied little between schools. Nevertheless, models were developed to explain the slight variation that did exist in the grade 12 gender gap. No school characteristics explained the variation in the gender gap in grades 4 and 8, while in grade 12, only one variable in the six models was able to explain any variation in the gap—the gender gap in grade 12 geometry was larger in schools where computers were used more frequently in math classes.

Results: predicting the race-ethnicity gap

Average race-ethnicity differences in math and geometry within schools were more pervasive than gender differences. In each grade, in both math and geometry, African-American, Hispanic, and Native American students averaged about 14 achievement points lower than did white and Asian-American students, controlling for gender, SES, and coursetaking. While this difference varied little among schools in grade 4, there was a wide variation in the size of the race-ethnicity gap in grades 8 and 12. A few variables explained this variation in the race-ethnicity gap. In grade 4 math and geometry, the race-ethnicity gap was smaller in schools where students spent more time on worksheets in math classes. In grade 8 math, working with objects was associated with a smaller race-ethnicity gap, whereas a smaller race-ethnicity gap in math in grade 12 was associated with a larger student/teacher ratio and more district instructional funds per student.

Lessons learned about using HLM with NAEP data

The outcomes of this study were the result of interactions between the type of data available in NAEP, the types of models that can be tested using HLM, and the way these interactions were expressed in the HLM estimates and statistics. Both NAEP and HLM worked very well when explaining variations in achievement. However, they did not work particularly well in explaining variations in the gender and race-ethnicity gaps. The reasons for this difference seem to stem from limitations of both NAEP and HLM. The HLM models were probably so successful in explaining the variations in the average NAEP achievement data because assessment scores, even in the form of plausible values, are the product of years of refinement by NAEP, and are the dataset's "best" variables. In addition, the intercept equation in HLM usually is the most reliable. In contrast, the gender and race-ethnicity gap equations are usually less reliable in HLM, and in this study they had little variance to explain. These equations are also more sensitive to problems with the predictor variables, and the NAEP data may not have been precise enough to capture the true associations between gender, race-ethnicity, and classroom and school climate factors.

Despite the lack of predictors of the gender and race-ethnicity gaps, NAEP and HLM should be used to continue to explore these gaps. Besides constructing and testing other predictor variables, variations of the gender and race-ethnicity gaps can be created by testing interaction terms or different combinations of gender and race-ethnicity. However, if even these simple models were not successful, it is possible that the limitations of the data and the model may prevent researchers using HLM models and NAEP data from obtaining more meaningful results in this area.

In the process of analyzing these results, two patterns in the results appeared that are also related to characteristics of both NAEP and HLM. The first pattern was a lack of a consistent relationship between the amount of parameter variance, reliability, significant variables, and the proportion of variance explained. However, this was only an apparent lack of consistency, because parts of the HLM models were broken up for presentation purposes. In fact, they are interconnected models, and each part of the model affects the other parts. The second pattern was a sensitivity of the models to slight changes in variables, which showed that variable specification and choice is very important.

Despite the problems of variable specification, lack of significant predictors, inconsistency in the statistics, and sensitive models, it is still possible to produce meaningful interpretations of HLM/NAEP results. First, the more the construction and univariate characteristics of a variable are known, the easier it is to explain their association with other variables. Second, variables that are significant across varying models with different control variables are more robust predictors than those that only appear in one model. A combined model can then be used to test the most theoretically important and predictive variables. Third, while statistical significance is necessary for interpreting a result, it may not always be sufficient. The practical significance of the results must be considered also. Given that the four anchor levels in math are 50 points apart, and the standard deviations around average achievement are between 30 and 40 points, variables that predict average achievement or gap differences of under 5 points may not be as important as those that predict changes of 10 or more points.

The questions of whether using more conditioning variables would make a difference in the results and whether the conditioning variables would perform better than the nonconditioning variables were only partially answered. While more results were obtained using the 1990 dataset than in the previous study using the 1986 dataset, most of these results were in the new models not tested in 1986, and the two comparable models actually had fewer results in 1990. Using conditioning variables in these models seems to have contributed to finding significant associations with average achievement, although it did not help explain the gender or race-ethnicity gap. Although most of the conditioning variables were student-level variables aggregated to the school level, many were significant and had moderate effect sizes. While conditioning variables were not always significant, their presence may have made it more likely to find significant results. This suggests that conditioning variables should be used whenever possible, even if they are aggregated to the school level from the student level. However, since most of the conditioning variables used in 1990 were not available or not tested in models in 1986, it was not possible to tell whether their significance in 1990 was actually due to the fact that they are conditioning variables.

Recommendations for NAEP

This study is only the beginning of a detailed analysis of the relationships between gender, race-ethnicity, and achievement using HLM and NAEP data. For example, only main effects were tested in this study in order to keep the interpretations as clear as possible. However, the next stage of this research could examine the interaction of race-ethnicity with both SES and gender using either interaction terms or separate models for subgroups. In addition, further research can investigate such areas as the associations between SES and attitudes towards math and between attitudes towards math and course-taking patterns. This study also identifies areas that can be further investigated using more qualitative research methods.

In many ways this study came up against the limits of the use of NAEP data, or any cross-sectional dataset, for studying school-level correlates of achievement and of gender and race-ethnicity differences in achievement. While many excellent indicators were included in NAEP, their presence led to a desire for more and even better measures. The following recommendations would improve the type of data that NAEP can provide for this type of research:

1) Since NAEP is a cross-sectional study, it is not possible to tell what type of achievement growth has occurred during an academic year, and thus, whether any associations between school factors and achievement are due to factors during that school year. However, if some measure of achievement ability at the beginning of school year was provided, the association between school factors and current achievement could be examined, controlling for previous achievement. This would still not be causal evidence of school effects, but it would refine the correlational results. Since HLM models are correlational, they cannot indicate any causal relationships. Therefore, appropriate caveats should be included in all HLM research that might have causal or evaluative implications.

2) Ideally, adequate samples of students within classrooms are needed to test associations between student achievement and classroom and teacher factors. Aggregating these factors across classrooms to the school level weakens the ability of these variables to explain student achievement.

3) Whether or not classroom samples are provided, other more refined classroom measures are needed, such as student and teacher interactions during math instruction and the gender and race-ethnicity composition of the classroom and math workgroups.

4) More information on the selection, derivation, validity, and reliability of the student background, classroom, and school climate questions would be helpful and would contribute to an understanding of what these variables actually measure. In addition, due to

large numbers of missing values in the teacher variables, either more information on the missing values or more successful teacher data collection is necessary.

5) Finally, the use of plausible values, particularly those that have been conditioned with student and school variables, is still somewhat mystifying to researchers using NAEP data to identify student and school correlates of achievement. It would be helpful to have a non-technical explanation of the use of conditioning variables to create proficiency scores and the justification of their use as subsequent predictors of those scores, including the procedure of aggregating student-level conditioning variables to the school level and the effect of using these as school-level predictors on the results.

Table of Contents

List of Tables

Chapter I

Background and Purpose

A. Introduction

This report illustrates the use of hierarchical linear models (HLM) with NAEP data to identify school and other correlates of student achievement. Based on an analysis of the 1990 NAEP mathematics achievement data for 4th, 8th, and 12th graders in public schools, this study is part of an ongoing exploratory effort to demonstrate the potential usefulness of HLM, a state-of-the-art statistical procedure with NAEP, a complex data set.

The purpose of this report is to illustrate several types of HLM analyses that can be performed on NAEP data. These demonstration analyses show how HLM and NAEP can be used to identify school and classroom correlates of student achievement, controlling for other school and student characteristics. In addition, the study also illustrates how the HLM methodology and NAEP data allow an examination of school-level correlates of achievement differences by gender and race-ethnicity within schools. These analyses are conducted not only on the overall subject of mathematics, but on geometry, one of the higher-level subscales of mathematics achievement. The theoretical models for the study are based on the literature on school effects and on research on the differences in math achievement by gender and race-ethnicity, while the methodology is guided by the emerging writings on HLM and by the author's past experience with HLM and NAEP.

This study builds upon and improves a multi-level analysis of school effects on math and science achievement that was conducted on the 1986 NAEP data.[1] That analysis was the first time HLM had been applied to the multi-level NAEP data, and it illustrated the value of using HLM when analyzing NAEP. This study differs from the Arnold, Kaufman, and Sedlacek analysis of the 1986 NAEP data in that it uses mathematics data only, rather than mathematics and science data, and several different models are developed. However, two of the 1990 models are similar to 1986 models, and the 1990 results from these models are compared to the 1986 results. In addition, this report differs from the earlier study report in that it focuses more on the methodology of using HLM with NAEP data than on the results of the study.

Due to the exploratory nature of the study and limitations of the data, policy changes are not recommended on the basis of this report. Instead, researchers are encouraged to use this analysis as the basis for an understanding of the procedures and questions involved in using NAEP data and hierarchical linear models for school effectiveness studies.

B. Research on school effects and school effectiveness

Research on "school effects" or "school effectiveness" seeks to identify the types of schools in which students attain higher achievement scores. Although the name implies causal relationships, most "school effects" studies, including this one, are actually "school

---

[1]C. Arnold, P. Kaufman, and D. Sedlacek, *School Effects on Educational Achievement in Mathematics and Science: 1985–86* (Washington, DC: U.S. Department of Education, National Center for Educational Statistics, 1992).

correlation with achievement" studies. School factors that have been found or theorized to be correlated with achievement include school fiscal and physical resources, student body and community characteristics, school social structure, school climate, and instructional organization and methods.[2]

Despite the large number of variables identified that are theoretically linked to student outcomes, school effects research has always found it difficult actually to explain much of the variance in student achievement, especially with physical and fiscal school-level factors.[3] Even within-school factors such as administrative and instructional organization and teacher characteristics and behavior have not been able to predict student learning adequately.[4] Centra and Potter suggested in 1980 that research done on factors closest to student outcomes, such as student characteristics and behavior, might produce more results, especially if teacher and school factors at all levels of influence were also taken into account.[5] By 1989, Oakes was also pointing to a more synergistic model among school resources, school structure, and school culture that could serve as indicators of effective schools if the factors were all somehow taken into account together.[6]

Of the types of school characteristics that were shown in the school effectiveness literature to be related to student achievement, six such groups of characteristics that could be measured in the NAEP dataset were identified and designated as separate models for this study. The groups were student body characteristics, fiscal, physical, and staff resources of the school, classroom instructional methods, and three aspects of school climate: attitudes towards math, student safety and behavior, and academic expectations. Using HLM allowed the relationship of these school-level characteristics to school and student achievement to be examined, while also taking into account the association of student-level characteristics with achievement.

C. Research on mathematics achievement differences by gender and race-ethnicity

Besides being interested in finding school predictors of student achievement overall, educators are also concerned about differences in achievement by gender and race-ethnicity within and between schools, and about the school and classroom factors that might be related to those differences. The concern is particularly strong in relation to math and science achievement because success in these subjects opens up opportunities in higher

---

[2]M. Rutter, B. Maughan, P. Mortimore, and J. Ouston, *Fifteen Thousand Hours: Secondary Schools and Their Effects on Children* (Cambridge, MA: Harvard University Press, 1979); W. B. Brookover, C. Beady, P. Flood, J. Schweitzer, and J. Wisenbaker, *School Social Systems and Student Achievement: Schools Can Make a Difference* (New York: Praeger, 1979); C. S. Anderson, "The Search for School Climate: A Review of the Research," *Review of Educational Research* 52 (3) (Fall 1982): 368–420; M. Rutter, "School Effects on Pupil Progress: Research Findings and Policy Implications," in L. Schulman and G. Sykes (eds.) *Handbook of Teaching and Policy* (New York: Longman, 1983): 3–41; T. L. Good and R. S. Weinstein, "Schools Make a Difference: Evidence, Criticisms, and New Directions," *American Psychologist* 41 (10) (1986): 1090–97; B. L. Wilson and T. B. Corcoran, *Successful Secondary Schools: Visions of Excellence in American Public Education* (London: Falmer Press, 1988); S. E. Mayer and C. Jencks, "Growing Up in Poor Neighborhoods: How Much Does It Matter?," *Science* 243 (March 1989).

[3]J. Oakes, "What Educational Indicators? The Case for Assessing the School Context," *Educational Evaluation and Policy Analysis* 11 (2) (Summer 1989): 181–99.

[4]J. A. Centra and D. A. Potter, "School and Teacher Effects: An Interrelational Model," *Review of Educational Research* 50 (2) (Summer 1980): 273–91.

[5]Ibid., 273–91.

[6]J. Oakes, "What Educational Indicators?"

paying fields. While females tend to stay in math and science courses until high school and early college, the overall achievement of females in math and science drops below that of males during high school. For African-American and Hispanic students, achievement in math and science falls short of whites from early on, and these groups tend to drop out of math and science before high school.[7]

In order to identify school and classroom factors that might be associated with smaller gaps in the achievement between males and females and between whites and African-American, Hispanic, and other race-ethnicity groups, more specific information is needed about these patterns. For instance, when the math and science achievement levels of females and African-Americans and Hispanics fall below those of males and whites and Asian-Americans, the differences are greatest in the higher order math and science skills.[8] For females, African-Americans, Hispanics, and other race-ethnicity groups, the relationships between lower achievement in these skills and the school and classroom characteristics that might mitigate these patterns in math and science need to be examined. Unfortunately, correlational analysis can only suggest associations between achievement and school and classroom characteristics. It remains for experimental research to confirm the presence and direction of any causal relationships.

Previous research on the relationship between math achievement and gender and race-ethnicity provides some guidance about variables to examine in this study. Peterson and Fennema found that gender differences in math achievement in elementary school may be related to type of classroom activities.[9] Their studies show that females and males performed differently depending on whether math activities were competitive or cooperative and whether the tasks were low- or high-level cognitive tasks. In an ethnographic study, Weis also found that in high school, higher math achievement of females may be associated with their perception of their need to work, which is based on their social class, positive attitudes toward their own work, and being on an academic track in high school.[10]

In addition, Fennema and Peterson hypothesized that working autonomously is a necessary way to learn high level math skills.[11] They developed an "autonomous learning behavior model," in which autonomous learning behaviors are based upon internal beliefs such as confidence in one's math abilities, the perceived usefulness of math, believing that learning math is congruent with one's gender, and a belief that success is due to one's own ability and effort. Autonomous learning is more likely to occur with traditional instructional methods such as separate seat work rather than in cooperative groups.[12]

---

[7]J. Oakes, *Lost Talent: The Underparticipation of Women, Minorities, and Disabled Persons in Science* (Santa Monica: The RAND Corporation, 1990): 15-20.

[8]Ina V. S. Mullis and Lynn B. Jenkins, *The Science Report Card: Elements of Risk and Recovery* (Princeton, NJ: Educational Testing Service, September 1988): 55-58 and John Dossey et al., *The Mathematics Report Card: Are We Measuring Up?* (Princeton, NJ: Educational Testing Service, June 1988): 54-58.

[9]P. L. Peterson and E. Fennema, "Effective Teaching, Student Engagement in Classroom Activities, and Sex-Related Differences in Learning Mathematics," *American Educational Research Journal* 22 (3) (1985): 309–35.

[10]L. Weis, "High School Girls in a De-Industrializing Economy," in L. Weis (ed.), *Class, Race, and Gender in American Education* (Albany, NY: SUNY, 1988): 183–208.

[11]E. Fennema and P. L. Peterson, "Autonomous Learning Behavior: A Possible Explanation of Gender-Related Differences in Mathematics," in L. C. Wilkinson and C. B. Marrett (eds.), *Gender-related Differences in Classroom Interactions* (New York: Academic Press, 1985): 17–35.

[12]For a review of this literature, see E. Fennema and G. C. Leder (eds.), *Mathematics and Gender* (New York: Teacher's College Press, 1990), especially Chapters 4 and 6.

Diverse factors have been associated with race-ethnicity differences in achievement. Ortiz identified several types of school and classroom resources that were not as available to Hispanic students as to non-Hispanic students, whether the Hispanic students were in bilingual or mainstream classrooms. These resources included material resources, such as books, computers, and school funding, and personal resources, which included number of teacher aides, credentials of teachers, teacher time and interactions with students, quality of teacher/student interactions, and the type of instructional method.[13] Oakes showed that minority groups who lack nonschool educational enrichment resources and activities such as books, newspapers, computers, and trips to museums and other cultural events also have lower achievement than whites and Asian-Americans, who are more likely to have these home background advantages.[14] In a summary of research on successful high school achievement, Oakes reported that overall achievement is based on "access to math and science instruction, early achievement in math and science," positive attitudes towards math and science, and "high expectations and encouragement" from surrounding adults.[15] In addition, she found that minority students, schools with higher percentages of minority and poor students, and students in nonacademic track classes receive lower level course content, types of thinking skills and topics, instructional methods, and homework expectations than other students and schools.[16] All of these factors are associated with lower achievement.

Ogbu points out that while African-American students may have very high educational aspirations and positive attitudes about learning, their actual behavior towards school work prevents them from reaching those goals.[17] He explains that the strengthening of African-American identity may involve rejecting characteristics associated with being white such as working hard in school, and that this phenomenon cuts across class lines.[18] Stanlaw and Peshkin report research that suggests that having minority students represent 15–40 percent of the total enrollment in a school supports racial harmony and positive identity for the minorities.[19] However, they cannot confirm that this is true for every minority group or school.

These studies suggest that the following variables might be associated with differences in achievement by gender or race-ethnicity: type of instruction; classroom activities and interactions; student confidence and perceived usefulness of math; attitudes about math and work; school, classroom, and teacher resources; expectations from adults; teacher and principal characteristics; high school program; and percentage of minority students in a school. Models expressing these concepts were tested using the variables available in the 1990 NAEP.

---

[13]F. I. Ortiz, "Hispanic-American Children's Experience in Classrooms: A Comparison between Hispanic and Non-Hispanic Children," in L. Weis (ed.), *Class, Race, and Gender in American Education* (Albany, NY: SUNY, 1988): 63–86.

[14]J. Oakes, "Tracking in Mathematics and Science Education: A Structural Contribution to Unequal Schooling," in L.Weis (ed.), *Class, Race, and Gender in American Education* (Albany, NY: SUNY, 1988): 106–25.

[15]J. Oakes, "Tracking in Mathematics and Science Education," 112.

[16]J. Oakes, "Tracking in Mathematics and Science Education," 117–18.

[17]J. O. Ogbu, "Class Stratification, Racial Stratification, and Schooling," in L. Weis (ed.), *Class, Race, and Gender in American Education* (Albany, NY: SUNY, 1988): 163–82.

[18]J. O. Ogbu, "Class Stratification, Racial Stratification, and Schooling," 163–82.

[19]J. Stanlaw and A. Peshkin, "Black Visibility in a Multi-ethnic High School," in L. Weis (ed.), *Class, Race, and Gender in American Education* (Albany, NY: SUNY, 1988): 209–29.

D. The use of HLM in school effects research

HLM is a multivariate regression-like technique that was developed specifically for use in school effects research. Before the development of HLM, school effects research had not been able to show conclusively that differences among schools were associated with different levels of student performance, in part because of persistent methodological problems. The major problem was that much of this research had not adequately modeled the multi-level nature of student achievement data. Students are nested within schools.[20] Thus, students exist at one level of analysis and schools exist at a higher level of analysis. Since student characteristics vary within schools and school characteristics vary between schools, questions about school effects and achievement require the simultaneous exploration of relationships at the within- and between-school levels. However, earlier school effects research relied primarily on single-level multiple regression models at either the student level or the school level to assess school effects, and therefore failed to model the multi-level structure of these relationships accurately. Treating these data as if they were all at the same unit of analysis may have led researchers to misleading conclusions about the effect (or noneffect) of various aspects of the school environment on student achievement.[21]

However, recent developments in the statistical theory of hierarchical linear models (HLM) solved these methodological problems. Some of the major theoretical and software applications of HLM were developed specifically to solve these problems in school effects research.[22] HLM allows direct representation of the association of school-level factors with student-level factors related to achievement within schools while controlling for the confounding factors at each level. It also partitions error variance into the appropriate level. In this way, it directly models the hierarchical nature of the data. As a result, within-school differences in achievement by student characteristics such as gender and race-ethnicity can be modeled as a function of school or classroom characteristics.[23]

Since its development, HLM has been applied to numerous school effects studies.[24] This research has been very successful at distinguishing between student-level and school-

[20] To be more exact, students are nested within classrooms within schools. An HLM analysis of these three levels is possible. However, there were not enough students per classroom or classrooms per school in the NAEP sample to analyze classroom differences as well as school differences. Therefore, this methodological discussion will focus on the student-level and school-level differences that were analyzed in this study.

[21] For an early warning on the dangers of using single-level models to model school effects, see L. Cronbach, *Research on Classrooms and Schools: Formulation of Questions, Design, and Analysis*, occasional paper of the Stanford Evaluation Consortium (Stanford, CA: Stanford University, 1976). For a review of the methods used before the advent of hierarchical linear models, see L. Burstein, "The Analysis of Multilevel Data in Educational Research and Evaluation," *Review of Research in Education* 8 (1980): 158–233.

[22] S. W. Raudenbush and A. S. Bryk, "A Hierarchical Model for Studying School Effects," *Sociology of Education* 59 (January 1986): 1–17; A. S. Bryk, S. W. Raudenbush, M. Seltzer, and R. Congdon, *An Introduction to HLM: Computer Program User's Guide*, 2nd ed. (Chicago IL: University of Chicago, Department of Education, 1988); A. S. Bryk and S. W. Randenbush, "Toward a More Appropriate Conceptualization of Research on School Effects: A Three-level Hierarchical Linear Model," in R. D. Bock (ed.), *Multilevel Analysis of Educational Data* (San Diego, CA: Academic Press, 1989); and A. S. Bryk and S. W. Raudenbush, *Hierarchical Linear Models for Social and Behavioral Research: Applications and Data Analysis Methods* (Newbury Park, CA: Sage, 1992).

[23] A. S. Bryk and S. W. Raudenbush, *Hierarchical Linear Models for Social and Behavioral Research: Applications and Data Analysis Methods* (Newbury Park, CA: Sage, 1992).

[24] See for example L. Bernstein, *The Application of Hierarchical Linear Modeling to Multilevel Student Achievement Data*, paper presented at the American Educational Research Association (Chicago, IL: 1991); V. E. Lee and A. S. Bryk, "A Multilevel Model of the Social Distribution of High School Achievement,"

level differences. Explaining these differences has proved more elusive, however, and researchers continue to posit models and to identify and specify variables that will account for more of the variance.

In several ways, school effects research is back at the beginning stages of a research program. On the one hand, school effects research seems to be starting over by testing the variables used in single-level models in new, multi-level models in order to determine which of the old theoretical models are still valid. On the other hand, new theoretical models are being developed based on the new questions that can be asked using HLM. Both efforts are needed, because at the same time that the methods have caught up to the theories, the school effects research and theory are beginning to catch up with the available methods. Recent research has confirmed the value of multi-level models, the importance of modeling variations in school effects on different subgroups of students, the need to take student body characteristics into account, and the importance of distinguishing between sampling and explainable (parameter) variance.[25] In addition, some researchers point out that longitudinal data is better than cross-sectional data in monitoring "the effects of school policies on changes in student performance."[26] However, as in single level models, both longitudinal and cross-sectional studies are valuable.

This cross-sectional study included variables based on past theoretical models of school effects that predict within-school achievement with between-school characteristics. In addition, new theoretical models were tested to predict within-school differences in achievement by gender and race-ethnicity based on between-school characteristics.

E. The use of NAEP in school effects research: strengths and limitations

This analysis relied on the main mathematics assessment data from the 1990 National Assessment of Educational Progress (NAEP). NAEP biennially tests a nationally representative sample of students in grades 4, 8, and 12 in public and private schools.[27] In addition, in 1990 NAEP also collected representative data from 37 states, the District of Columbia, and two territories, in the first trial of a series of voluntary state-level

---

Sociology of Education 62 (1989): 172–92; S. W. Raudenbush and A. S. Bryk, "A Hierarchical Model for Studying School Effects," *Sociology of Education* 59 (January 1986): 1–17; R. W. Rumberger and J. D. Willms, *The Impact of Racial and Ethnic Segregation on the Achievement Gap in California High Schools*, paper prepared for the Annual Meeting of the American Educational Research Association (Chicago, IL: 1991); J. D. Willms and M. Chen, "The Effects of Ability Grouping on the Ethnic Achievement Gap in Israeli Primary Schools," *American Journal of Education* 97 (3) (1989): 237–57; and J. D. Willms and S. Jacobsen, "Growth in Mathematics Skills During the Intermediate Years: Sex Differences in School Effects," *International Journal of Educational Research* 14 (1990): 157–74. For a review of educational applications, see S. W. Raudenbush, "Educational Applications of Hierarchical Linear Models: A Review," *Journal of Educational Statistics* 13 (2) (1988): 85–116. For a recent collection of applications, see S. W. Raudenbush and J. D. Willms, *Schools, Classrooms and Pupils: International Studies of Schooling from a Multilevel Perspective* (San Diego, CA: Academic Press, 1991).

[25] J. D. Willms and S. W. Raudenbush, "A Longitudinal Hierarchical Linear Model for Estimating School Effects and Their Stability," *Journal of Educational Measurement* 26 (3) (1989): 209–32 and J. D. Willms, *Monitoring School Performance: A Guide for Educators* (London: Falmer Press, 1992).

[26] J. D. Willms and S. W. Raudenbush, "A Longitudinal Hierarchical Linear Model for Estimating School Effects and Their Stability," *Journal of Educational Measurement* 26 (3) (1989): 209–32.

[27] Students in grades 4, 8, and 12 were sampled starting in 1988. Between 1969–70 and 1986, most NAEP samples consisted of students ages 9, 13, and 17. In 1986, students in grades 3, 7, and 11 were also sampled. See U.S. Department of Education, National Center for Education Statistics, *The NAEP 1990 Technical Report* (Washington, DC: Government Documents, February 1992).

assessments.[28] The primary goals of NAEP are to detect and report the current status of, as well as changes in, the educational attainments of American students. NAEP meets these goals by thoroughly testing students on a wide variety of items and subscales in reading, writing, math, science, and other subjects. However, besides achievement data, NAEP also collects a rich array of contextual information about the students, their teachers, and their schools.

This comprehensive set of data makes NAEP well-suited for a study of student, classroom, and school correlates with student achievement. Many of the contextual variables in NAEP are drawn from the school effects literature, so using them allows the testing of theoretical relationships between variables and achievement. In addition, NAEP includes math attitude questions that are based on standard scales used in research on gender and math, including a specific question about gender and math.[29] There are also numerous other variables that can be used to explore the relationships between achievement, gender, and race-ethnicity.

As a cross-sectional data set, NAEP provides the most complete "snapshot" of American elementary and secondary student achievement currently available. The detailed achievement data together with the collection of information at the various levels of student, classroom, teacher, principal, and school make it unique among national datasets. In addition, while the three grades included—4th, 8th, and 12th—provide a comparison of students in elementary school, middle school, and high school at the same point in time, the biennial fielding of NAEP allows comparisons over time.

However, NAEP has limitations related to sample design, variable selection and construction, data collection, and proficiency score estimation that raise questions about the validity and interpretation of the results in the areas of classroom instruction and school climate, especially in complex studies such as this. Four concerns that could affect the interpretations of this study are reviewed here.[30]

First, the sample design of NAEP requires collecting data on about 20-40 students in each school across classrooms in each grade, and data are available about the students, about instruction in their classrooms (from both students and teachers), and about the schools. However, the classroom data cannot be identified by classroom, but can only be connected to each student, or aggregated across the school. Even if the data could be grouped by classroom using the teacher ID, there would be too few students per classroom to analyze. Therefore, the estimation of any relationship between classroom instruction and average school achievement, which is what this study attempts to do, is diluted by the aggregation of instructional patterns from a variety of classrooms.

While these classroom data could be used within schools to analyze the relationship across students, the fact that an unknown number of students would be in the same classrooms would bias the variance. Ideally, enough classroom data would be available to estimate a three-level HLM model of students within classrooms within schools, but that would greatly expand the NAEP sampling design. Meanwhile, the classroom instructional variables, whether they come from the teachers or students, lose much of their strength in

---

[28]U.S. Department of Education, National Center for Education Statistics, *NAEP 1990 Trial State Assessment Secondary-use Data Files User Guide* (Washington, DC: Government Documents, June 1991).

[29]E. Fennema and J. Sherman, "Fennema-Sherman Math Attitudes Scales," *JSAS: Catalog of Selected Documents in Psychology* 6 (1) (Ms. No. 1225, 1979): 31.

[30]This discussion is based, in part, on comments by L. Burstein about another NAEP report, documented in a letter to S. Shakrani, Design and Analysis Branch, Education Assessment Division, NCES, March 8, 1994.

their aggregation to the school level, and it would not be surprising if they produced few significant results.

Second, while the many of the student, classroom, and school climate questionnaire items seem based on school effects, gender, and race-ethnicity research, there is little information in the NAEP technical manual about how these items were selected or derived, and whether they have been validated by NAEP or previous questionnaire developers.

For instance, the validity and reliability of measuring types of math instructional methods by gross frequency rather than by minutes a day was not provided, and the ability of students to estimate either the frequency or minutes of instruction was not mentioned. While it was helpful to have the student estimates of these variables to back up the teacher information (see below), the reliability of their perceptions remained a question. Likewise, the measurement of the home environment with only reading materials seems to need justification. At least one item, on math and gender, is the only question on that topic, and since there are a wide variety of ways in the literature to measure that attitude, the rationale for using this particular question is unclear. In addition, this item is included in a NAEP-derived variable about positive perceptions about mathematics, although it is debatable whether it measures the same dimension of perceptions. Overall, in order to build complex models, the variables used in those models must be reliable indicators of underlying phenomena, and with no information about these items, doubts do arise.

Third, one part of the sample design is to collect data from the teachers of some of the students in grade 4 and grade 8 math classes. While this is better than no information on teachers, teacher and classroom data would be more powerful if there were classrooms full of students to go with them, as mentioned above. However, given the current sample design, an even greater problem exists with the teacher data—the number of missing values that made most of the teacher data unusable. Some of the missing values may be due to the fact that not all of these students were taking math. However, it was not clear whether too few teachers had been sampled, too few students were taking math, or too few teachers had responded. For whatever reason, there were too many missing values on teacher variables, so the result was an inability to use the teacher characteristics or instructional variables.

Finally, providing five plausible values and rather just one score for student proficiency makes NAEP a more challenging dataset than others with more traditional measures of achievement. This report and the software developed during this study are designed to demystify and simplify the use of plausible values and their statistics. However, these procedures, and the lack of a non-technical explanation of their necessity, may still be a barrier to the wider use of NAEP.

F. Assumptions of causality in NAEP HLM studies

NAEP provides cross-sectional, correlational data rather than a longitudinal or experimental design. While it is tempting in an HLM school effects study to make causal assumptions and conclusions, it is important to remember that the correlational design of the study allows only the identification of associations rather than causal relationships between school and student factors and student achievement.

In addition, although NAEP is cross-sectional, the school, teacher, and student variables in NAEP do not necessarily occur at one point in time, or at any known time for a known duration, so it is not possible to know whether any of the school factors could have preceded achievement and thus influenced student achievement. First, although the assessment occurs in the Spring, there is no measure of a student's previous achievement score, so it is not possible to determine the level of achievement that has been gained during that school year. In addition, there is no way to know whether the school factors, teacher characteristics, or classroom instructional methods have been in effect all year, or have been constant for years.

Similarly, there is no way to know whether students have been in their respective schools long enough for that school to have had an impact on their achievement. It is more likely that the students in grades 4 and 12 rather than grade 8 would have been in the same school during the previous year or so, because of the grades included in elementary and high schools. However, that would not be true if they recently moved. The students in grade 8 would have been in the same school in previous years only if they attended an elementary school that included 8th grade or a middle or combined school that started earlier than 8th grade. If they were in their first year of a new middle or high school, or if they had recently moved, they might not have been in the school long enough for achievement to be affected.

In order to prevent any assumptions of causality, this report has avoided the use of the words with cause-and-effect connotations such as "effect." Instead, the association between variables with significant coefficients has been emphasized. Any causal implications remaining are unintentional, and need to be confirmed with more experimental research.

G. Past HLM analysis of NAEP data

Arnold, Kaufman, and Sedlacek performed the first HLM analysis of NAEP data, and solved many logistical and statistical problems in using both together.[31] That study, which used 1986 NAEP data, modeled traditional school effects variables as correlates to student achievement in math and science in three grades. However, those models resulted in few significant school or teacher characteristics. One explanation was that the models posited to explain differences in achievement by gender and race-ethnicity were based on school effects research, but not on gender and race-ethnicity research.

Another explanation for this lack of results was the low number of "conditioning" variables used to create the 1986 NAEP plausible values of achievement scores. Conditioning variables are those student, teacher, and school-level variables that were used to estimate the student plausible values of the achievement scores.[32] If variables not used to impute the plausible variables are used in regression models, their coefficients are mis-estimated.[33] Since few conditioning variables were used to create the 1986 plausible values, this limited the number of variables that could produce reliable and unbiased regression estimates.

---

[31]C. L. Arnold, P. D. Kaufman, and D. S. Sedlacek, *School Effects on the Relationship Between Science Achievement and Gender, Race-ethnicity, and SES in Grades Three, Seven, and Eleven: 1985-86*, paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL: 1991); and C. L. Arnold, P. D. Kaufman, and D. S. Sedlacek, *School Effects on Educational Achievement in Mathematics and Science: 1985–86* (Washington, DC: U.S. Department of Education, National Center for Education Statistics, 1992).

[32]For more information about conditioning variables, see chapter two. For more information on the creation and use of plausible values, see A. Rogers et al., *National Assessment of Educational Progress: 1990 Secondary-use Data Files User Guide* (Princeton, NJ: Educational Testing Service, March 1992).

[33]R. J. Mislevy, *Randomization-Based Inferences About Latent Variables From Complex Samples* (Princeton, NJ: Educational Testing Service, September 1988).

These problems were addressed in the current study. First, the 1990 NAEP used many more conditioning variables than the 1986 NAEP to create the plausible values, and these conditioning variables were used whenever possible as predictors of student achievement. Second, several models were developed specifically to predict average achievement differences by gender and race-ethnicity, based on research on gender and race-ethnicity differences in achievement. Third, besides predicting math achievement overall, the geometry subscale was also used as an outcome variable in order to investigate gender and race differences in higher order math skills.

H. Purpose and organization of report

The purpose of this report is to demonstrate the type of HLM analyses that can be performed on NAEP data. For this purpose, a research study was conducted using HLM on NAEP data, and the results of that study are discussed. The study is presented in the context of the methodology, and the results of the study are used to illustrate aspects of the methodology. However, in order to provide continuity with earlier research, the results of this study are compared to an earlier, similar study.

The first and major chapter discusses the methodology of the study. This chapter starts with an introduction to the statistical method of HLM, and describes the process of developing HLM models, including special considerations needed when using NAEP data. Then, the data sources, sample, and the choice and construction of models and variables used in this analysis are discussed. Finally, the technical details of running HLM on NAEP data and the interpretation of HLM statistics are explained.

The following three chapters present the results of the analysis for 1) models explaining school differences in average achievement, 2) models explaining school differences in achievement by gender, and 3) models explaining school differences in achievement by race-ethnicity. Results are presented in text tables and discussed in the text. Appendix A contains supporting tables of the HLM results.

Following the presentation of the results of the models, the next chapter contains the results of two exploratory analyses investigating alternative models for explaining differences by gender and race-ethnicity. These analyses are reported and discussed. The report ends with a summary and discussion of the results, a discussion of what was learned about using HLM with NAEP, and a set of recommendations for further research, the use of HLM, and changes in NAEP.

Chapter II

Methodology

This chapter presents the methodology of developing and estimating HLM models, using the current study as an example. The first section provides a brief introduction to hierarchical linear models (HLM) and discusses the types of models that can be tested using the basic two-level HLM. The second section describes the process of developing HLM models, including special considerations needed when using NAEP data. This section also describes the data preparation and variable creation for the HLM models used in this study. The third section details the software logistics of estimating HLM models with NAEP data. The final section discusses the statistics used to interpret HLM results, including the interpretations of these statistics in this exploratory study.

A. Hierarchical Linear Models (HLM)

Overview of hierarchical linear models (HLM) [34]

Like most data about schools and student achievement, the data collected under NAEP is hierarchical in nature because students, at one level of analysis, are nested within schools, at the next higher level of analysis. Hierarchical linear models address the problem of students nested within schools in the following way. Using a sample of schools with a sample of students in a particular grade within each school, a student-level linear regression model is estimated for each school to predict the association of student characteristics with student achievement in that grade. This is the level-1 equation.

Simultaneously, a school-level regression model is estimated for the schools at the school level to predict the association of school characteristics with each of the school-level estimates—the intercept and each coefficient—from the student-level models. This is the level-2 equation. Separate estimates are produced for the variance at level one within schools and the variance at level two between schools. Conceptually, HLM consists of estimating regressions of regression results, except that the equations at each level are estimated at the same time rather than sequentially, and the variance at one level is taken into account in estimating the next level.

[34] This overview is based on A. S. Bryk and S. W. Raudenbush, *Hierarchical Linear Models: Applications and Data Analysis Methods* (Newbury Park, CA: Sage, 1992).

Two-level HLM equations

Each two-level HLM analysis consists of the following steps.[35] In the first step, the within-school models are estimated using ordinary least squares (OLS) regression analysis. For instance in this study, the most basic within-school model estimates achievement as a function of the following student characteristics—gender, race-ethnicity, and SES. This results in an equation for each grade level in each school that consists of regression coefficients (called *Betas* in HLM) that estimate the association of achievement with being female, with race-ethnicity, and with SES level for student in that grade level in that school. The equation also estimates an intercept, which represents the average achievement in the school. Within each school, the equation at each grade level takes the form of the following regression equation:

Within-school student-level equation[36]

$$y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + r_{ij} \qquad (2.1)$$

where:

i represents the i^{th} student
j represents the j^{th} school
y_{ij} represents the achievement score of the i^{th} student in the j^{th} school
β_{0j} is the intercept, or the average achievement in the j^{th} school
β_{1j} is the Beta coefficient for gender in the j^{th} school
β_{2j} is the Beta coefficient for race-ethnicity in the j^{th} school
β_{3j} is the Beta coefficient for SES in the j^{th} school
X_{1ij} represents the gender of the i^{th} student in the j^{th} school
X_{2ij} represents race-ethnicity of the i^{th} student in the j^{th} school
X_{3ij} represents SES of the i^{th} student in the j^{th} school
r_{ij} is random error in the j^{th} school.

---

[35]These steps are actually simultaneous, but they can be understood most easily as sequential.

[36]The forms of these equations are taken from A. S. Bryk and S. W, Raudenbush, ***Hierarchical Linear Models: Applications and Data Analysis Methods*** (Newbury Park, CA: Sage, 1992), and C. L. Arnold, "An Introduction to Hierarchical Linear Models," ***Measurement and Evaluation in Counseling and Development*** 25 (2) (July 1992): 58–90.

In the second step of the HLM analysis, the intercept and the regression coefficients from the first step in the analysis become the outcome measures in the second step. That is, each of these within-school parameters—the intercept and the other Betas—is used as a dependent variable in a separate equation and the variation in these within-school parameters is modeled. These between-school equations produce coefficients (called *Gammas* in HLM) that estimate the association of each school-level characteristic with (in the case of this study) either the average achievement, the differences in achievement by gender, or the differences in achievement by race-ethnicity within the schools.

When the school-level equations are estimated in HLM, the data from each school are weighted by the inverse of their variance around these parameters. That is, the sohools with the most variance (usually those from smaller samples) are given less weight in contributing to the final school-level parameter estimates.

At the within-school level, HLM requires researchers to specify which within-school variables will be modeled with random parameter variance and which will be specified with fixed parameter variance. If the variable is considered to be random, then the parameter variance around its parameter coefficient is expected to vary randomly between schools. If the variable is considered to be fixed, then its parameter coefficient is expected to be the same in each school and its parameter variance is set to 0. The usual purpose of fixing is to allow a more efficient estimate of HLM models if there is in fact no variation around the parameter. However, another purpose of fixing is to add control variables to a within-school equation without losing degrees of freedom in the estimation. For a more detailed discussion of the use of variables with fixed and random parameter variance, see the exploratory analyses in chapter six.

In this study, the intercept, gender, and race-ethnicity parameters were allowed to vary randomly, and this variation was modeled as a function of the school-level characteristics across schools. The SES parameter coefficient was assumed to be the same for all schools, so SES was used as a control variable, and its variation was fixed and not modeled.

The following equations illustrate these models. First, the unconditional models are shown. These are the school-level models with only their intercept. They are called unconditional because they are not conditioned on, or predicted by, any school characteristics. Following the unconditional models are the conditional models, which include the school characteristics as predictors for the random parameters.[37] If the equation includes the random error term, u_{pj}, then the parameter variance has been designated as random. If there is no random error term, the parameter variance has been designated as fixed, i.e., Var (u_{pj}) =0, so the parameter, β_{pj} is assumed to be the same for all j schools.

Between-school school-level equations

a) Unconditional (before any school characteristics are added as predictors)

$\beta_{0j} = \gamma_{00} + u_{0j}$ (Intercept equation) (2.2)

$\beta_{1j} = \gamma_{10} + u_{1j}$ (Gender gap equation) (2.3)

$\beta_{2j} = \gamma_{20} + u_{2j}$ (Race-ethnicity gap equation) (2.4)

$\beta_{3j} = \gamma_{30}$ (SES equation) (2.5)

where:

β_{0j} represents the intercept, or the average achievement in the j^{th} school

β_{1j} represents the gender coefficient in the j^{th} school

β_{2j} represents the race-ethnicity coefficient in the j^{th} school

β_{3j} represents the SES coefficient in the j^{th} school

p is the number of within-school parameter equations (from 0 to 3 in this example)

γ_{p0} is the intercept, or the average within-school parameter value in the p^{th} equation

u_{pj} is random error in the p^{th} equation.

[37]Three similar terms are used in this report. "Conditional" and "unconditional" HLM models refer to the level-2 models with and without any independent variables. "Y conditoned on X" refers to independent variables (X) used to predict the dependent variable (Y) in the conditional models. "Conditioning" variables refer to variables used to create NAEP plausible values. All three terms are slightly different labels for independent variables used in regression equations.

b) Conditional (with school characteristics added as predictors)

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_{01j} + \gamma_{02}W_{02j} + \ldots + \gamma_{0m}W_{0mj} + u_{0j} \tag{2.6}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_{11j} + \gamma_{12}W_{12j} + \ldots + \gamma_{1m}W_{1mj} + u_{1j} \tag{2.7}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}W_{21j} + \gamma_{22}W_{22j} + \ldots + \gamma_{2m}W_{2mj} + u_{2j} \tag{2.8}$$

$$\beta_{3j} = \gamma_{30} \tag{2.9}$$

where:

γ_{p1} is the Gamma coefficient for the first school-level variable in the p^{th} equation
γ_{p2} is the Gamma coefficient for the second school-level variable in the p^{th} equation
γ_{pm} is the Gamma coefficient for the m^{th} school-level variable in the p^{th} equation
W_{p1j} represents the value of the first school-level variable in the j^{th} school in the p^{th} equation
W_{p2j} represents the value of the second school-level variable in the j^{th} school in the p^{th} equation
W_{pmj} represents the value of the m^{th} school-level variable in the j^{th} school in the p^{th} equation
m is the number of school-level parameter variables.

Questions that can be asked with HLM models

By designating the intercept and coefficients from the within-school equations as the dependent variables in the school-level equations and correctly modeling within- and between-school variation in achievement, HLM directly models the hierarchical nature of the data. These HLM models allow us to explain average student achievement in a school (β_{0j}, the intercept in the within-school equation) as a function of school characteristics. In addition, HLM allows us to explain the association between student characteristics, such as gender or race-ethnicity, and achievement within schools (the β_{pj} coefficients from the within-school equation) as a function of school characteristics. The γ_{pm} coefficients, or Gammas, from the between-school equations are the major indicators of school correlates with average achievement and of school correlates with the association of gender and race-ethnicity with achievement.

These equations allow us to ask several types of questions. The intercept equation (Eq. 2.6) measures the association of school characteristics, such as average amount of time spent taking math tests, with the average achievement in schools. This part of the model addresses questions such as, did schools with more than the average amount of time spent taking math tests have higher average achievement levels than schools with less than the average amount of time spent taking math tests?

The gender parameter equation (Eq. 2.7) measures the association of school characteristics, such as the average amount of time spent in small groups, with the gap in achievement between females and males, a gap that varied between schools. It addresses the question, did schools where students worked in small groups more than the average amount have a smaller or larger gap in achievement between females and males compared to schools where students worked in small groups less than the average amount?

The race-ethnicity parameter equation (Eq. 2.8), measures the association of school characteristics, such as the average amount of time spent in small groups, with the gap in achievement between African Americans/Hispanics/Native Americans and whites/Asians, a gap that varied between schools. It addresses the similar question, was the school-level characteristic of time spent in small groups associated with a smaller or larger gap between African Americans/Hispanics/Native Americans and whites/Asians?

Therefore, the types of questions researchers can use HLM models to answer are those which ask either how school characteristics are associated with average achievement in schools, or how school characteristics are associated with such student characteristics as the gender or race-ethnicity gap in schools. These are exactly the types of questions that school effects researchers and gender and race-ethnicity researchers ask, but until the advent of HLM they had no good statistical way to answer.

B. Selecting and Preparing NAEP data for HLM Models

This section contains explanations of the sample selection, the considerations involved in developing HLM models and choosing particular variables for an HLM analysis with NAEP data, the variable creation process, and the data and variable preparation process.

Choosing the NAEP sample

The study in this report uses the 1990 National Assessment of Educational Progress (NAEP) in mathematics to examine the relationship between school-level data and individual student-level math test data for a nationally representative sample of 4th, 8th, and 12th graders in public schools. This particular sample was selected for the following reasons. Mathematics, especially higher-level mathematics, is an important policy area for educational equity by gender and race-ethnicity. By focusing only on mathematics, not only could school correlates with overall mathematics scores be examined, but school correlates with subscale scores for geometry, a higher-level mathematics area, could be examined. Public schools were chosen both in order to control for type of school, and because policy changes could only directly affect public schools.

The three grades—fourth, eighth, and twelfth—were of interest in order to examine these correlates for students in elementary, middle, and high school, since different gender and race-ethnicity dynamics operate at these different levels. NAEP tests students in these three grades, regardless of age. However, NAEP also tests students in appropriate ages, regardless of grade. Therefore, choosing students in the target grades eliminated students in other grades. Only the target grades were chosen because mathematics achievement levels are seen as a function of grades, rather than ages.

The study sample was obtained in the following way from the NAEP main mathematics assessment files. The NAEP files consist of student files of all students assessed in a particular subject and grade/age group. In these files are the student-level proficiency scores as well as the student background data. These student files from the math assessment also contain data from the math teachers of 4th graders and 8th graders who were taking 4th grade or 8th grade math. These teacher data are connected to the particular student. NAEP also provides school files of the schools that participated in each assessment, for each subject and grade/age group. These files consist of data about the school and its community obtained from NAEP and from a school characteristics, principal characteristics, and school policies questionnaire that was distributed to each sampled school. About 93 percent of the schools completed these questionnaires.[38]

In the main mathematics assessment student files, there were 8,790 students in grade 4/age 9, 8,634 students in grade 8/age 13, and 8,406 students in grade 12/age 17. In the main mathematics assessment school files, there were 815 grade 4/age 9 schools, 688 grade 8/age 13 schools, and 596 grade 12/age 17 schools.[39]

---

[38]A. Rogers et al., *National Assessment of Educational Progress: 1990 Secondary-use Data Files User Guide* (Princeton, NJ: Educational Testing Service, March 1992).

[39]A. Rogers et al., *National Assessment of Educational Progress: 1990 Secondary-use Data Files User Guide* (Princeton, NJ: Educational Testing Service, Revised-June 1992) and U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th, 8th, and 12th Grade Students, Restricted-Use Data Base.

For this study, only students in the target grades were selected from the student files, and this sample consisted of 6,467 students in grade 4, 6,473 students in grade 8, and 6,311 students in grade 12. From the school files, only public schools were selected, creating a sample of 557 grade 4/age 9 schools, 424 grade 8/age 13 schools, and 458 grade 12/age 17 schools. When these files were merged, so that only students in target grades who were also in public schools were included, the final sample consisted of 5,080 students in 257 schools in grade 4, 5,198 students in 174 schools in grade 8, and 4,953 students in 186 schools in grade 12. In this sample, the grade 4 schools averaged 20 students per school, with a range of 4 to 46 students/school, the grade 8 schools averaged 30 students per school, with a range of 1 to 63 students/school, and the grade 12 schools averaged 27 students per school, with a range of 4 to 49 students/school.

Deleting private schools and choosing only target grades for this sample created an unsystematic subsample from the original mathematics assessment student and school samples. This violated the sample design and weighting scheme of the total math sample. Since the weights meant for the total math sample were used, the estimates in this sample will be somewhat biased.

Developing HLM Models from NAEP Data

In developing the HLM models in this study, there were five different goals or considerations to take into account when constructing models and choosing variables. The first goal was to replicate models from the earlier HLM study on the 1986 NAEP data. These models had all been based on the school effects literature. The second goal was to test new models using new or different NAEP variables, specifically those hypothesized to relate to differences in achievement by gender and race-ethnicity. A third goal was to use as many "conditioning" variables as possible, since the 1990 NAEP dataset contained more conditioning variables than the 1986 NAEP dataset. A fourth consideration was to choose which within-school variables would be random and which would be fixed, and to choose the appropriate school-level variables to model the random student-level parameters. A fifth consideration was to decide whether to estimate several separate models or to estimate one general model. After these considerations have been taken into account, then one can proceed to determine the appropriateness and availability of particular variables in the NAEP dataset, including checking for collinearity.

These five goals and considerations each influenced the final choice of models and variables. In this section, the background, importance, and decision of each goal or consideration is reviewed. In the following section, the final models and variables are presented.

1) Duplicating earlier models

One goal was to replicate the models tested in the earlier study with the 1990 NAEP data. However, there were some changes in variables between the 1986 NAEP and the 1990 NAEP datasets, and the 1990 NAEP had more missing values than the 1986 NAEP, especially in the areas of school resources and teacher and principal characteristics. Therefore, only two out of six models used on the 1986 NAEP data were replicated, and even in these models, the variables were not exactly the same. These models—the student body characteristics and the school resources—are discussed in the next section.

2) Developing new models

Another goal was to test new models using new or different NAEP variables, specifically those related to differences in achievement by gender and race-ethnicity. These new models focused more on the classroom instruction and school climate factors that are believed to be more related to student achievement and to gender and race-ethnicity differences in achievement than are physical and fiscal school characteristics. These models are discussed in the next section.

3) Maximizing the use of conditioning variables

Conditioning variables are related to plausible values, which are the product of the NAEP assessment procedures. Like previous NAEP Assessments, the 1990 Mathematics Assessment employed a variant of matrix sampling called balanced incomplete block (BIB) spiraling. With this procedure, the total assessment battery is divided into several 15-minute blocks of test items and two 5-minute blocks: one of student background characteristic items and one of content area items about math instruction, coursetaking, and attitudes. The background and content blocks were common to all students at each grade level. Each student was administered a booklet containing three blocks of test items as well as the two 5-minute blocks of background and content items. The BIB part of the method assigns blocks of test items to booklets in such a way that each pair of blocks appears in at least one booklet. This generates a number of different booklets. The spiraling part of the method then cycles the booklets for administration, so typically no two students in any assessment session in a school, and at most only a few students in schools with multiple sessions, receive the same booklet. At each age/grade level in the main math assessment, there were seven booklets. Each block of items was administered to approximately 3,700 students and each pair of blocks to approximately 1,225 students.[40]

Item response theory (IRT) was then used to estimate proficiency scores for each individual student, since each students had been tested on only a portion of the test items. However, these proficiency scores are latent variables conditional on the student's responses to several cognitive and background items and *are not directly observed*. That is, a set of cognitive and background variables (referred to as *conditioning variables*) as well as the student's assessment item reponses predicted the proficiency score a similar student would earn if they had answered all the items. Because the proficiency scores are not observed but estimated, there is some amount of uncertainty or variance associated with them. Thus, rather than having a single observed achievement score, there is a range or distribution of plausible values for each sampled student's proficiency. Each plausible value for an individual is not a scale score for that individual but may be regarded as a representative value from the distribution of potential scale scores for all students in the population with similar characteristics and identical patterns of item response.

In the NAEP dataset there are five such plausible values for each sampled student resulting from five random draws from the conditional distribution of plausible values for each student. Section C in this chapter explains how to use these five plausible values in the HLM analysis of achievement.

Conditioning variables are thus those student, teacher, and school-level variables that were used to estimate the student plausible values of the achievement scores. One goal of this analysis was to maximize the use of conditioning variables in the estimated models. In

[40]A. Rogers et al., *National Assessment of Educational Progress: 1990 Secondary-use Data Files User Guide* (Princeton, NJ: Educational Testing Service, March 1992).

the 1986 study, very few school or teacher characteristics were found to be significantly related to achievement. One explanation for this lack of results was the low number of conditioning variables used to create the plausible values of achievement scores during that year.

Studies by the Educational Testing Service have shown that statistics that involve variables that were included in the imputation of the plausible values for student proficiency scores are consistent estimators of population values. However, statistics involving background variables that were not used in the imputation of the plausible values have been shown to be biased. In particular, analyses of reading proficiency scores in the 1984 NAEP Reading Assessment indicated that multiple regression coefficients for nonconditioning variables tend to be underestimated by an average of 30 percent.[41] However, while misestimating the effects of nonconditioned variables, the direction of effects of nonconditioning variables are almost always correct. Therefore, one goal when using regression-like analysis with NAEP is to use as many of these conditioning variables as possible.

The 1990 study was designed to determine whether using more conditioning variables would make a difference in the results and whether the conditioning variables would perform better than the nonconditioning variables. The 1990 NAEP used many more conditioning variables than the 1986 NAEP. Although many of these conditioning variables had large numbers of missing values, both their missing and nonmissing values were used to estimate plausible values. Therefore, HLM analyses of the 1990 NAEP had a greater likelihood of producing significant results if models contained as many of the conditioning variables as possible. In developing the models for this analysis, conditioning variables were used wherever possible. In this analysis, most of the school-level variables used in the composite variables created in this analysis are conditioning variables, that is, they were used in the imputation of the plausible values, although some are not.

However, many teacher-level conditioning variables had too many missing values to use in the models. In order to increase the use of conditioning variables, several student-level conditioning variables were aggregated to the school level, where it is believed that they still "counted" as conditioning variables. Also, variables that were necessary to the model but could not be represented by conditioning variables were used. For these latter variables, although the analysis has correctly informed us on the direction of their effects, the size of these effects may have been underestimated by some unknown amount. In this discussion of HLM models as well as in the final discussion in chapter six, variables are labeled as either conditioning or nonconditioning variables.

4) Choosing between random and fixed-slope models

At the within-school level, HLM requires researchers to specify which within-school variables will be modeled with random parameter variance and which will be specified with fixed parameter variance. If the variable is considered to be random, then the parameter coefficient is expected to vary randomly between schools. If the variable is considered to be fixed, then its parameter coefficient is expected to be the same in each school and the parameter variance is set to 0. In this case, the coefficient can only vary nonrandomly based on the variable's actual variation in the sample of schools. The usual purpose of fixing is to allow a more efficient estimate of HLM models if there is in fact no variation around the parameter. In this case, the nonrandom variation of this parameter is not usually modeled,

---

[41] R. J. Mislevy, *Randomization-Based Inferences about Latent Variables from Complex Samples* (Princeton, NJ: Educational Testing Service, September 1988).

although it can be. However, another purpose of fixing is to add control variables to a within-school equation without losing degrees of freedom in the estimation.

In this study, the within-school variables of interest, gender and race-ethnicity, were allowed to vary randomly and were modeled, while the control variables of SES and coursetaking were fixed and were not modeled. The intercept, gender, and race-ethnicity parameters were allowed to vary randomly because they were the variables of interest. Their variation was modeled as a function of the school-level characteristics across schools. The between-school equations produced the *Gamma* coefficients that estimated the association of each school-level characteristic with either the average achievement in schools, the association of gender with achievement, or the association of race-ethnicity with achievement within the schools.

The SES and course-taking parameters were fixed, both because their assocation with achievement was not expected to vary between schools, and because they were primarily being included as control variables. Fixing control varibles saves degrees of freedom in the within-school models. For a more extensive discussion on how to determine statistically whether variables should have a fixed and random parameter variance, see the exploratory analyses in chapter six.

5) Deciding whether to estimate separate or combined models

As in the earlier study, this study developed and tested six separate between-school models, rather than testing one general model of all the variables or testing six models and combining the significant variables into one final model. These six separate models reflected the six groups of school characteristics that were deemed to be of theoretical importance based on previous school effects research—student body characteristics, fiscal, physical, and staff resources of the school, classroom instructional methods; and three aspects of school climate: attitudes towards math, student safety and behavior, and academic expectations. However, student body characteristics were included as control variables in each model.

These models were tested separately for the following theoretical and practical reasons. First, dividing the variables into six models avoided over-controlling with too many variables and obscuring some effects that might be significant. Secondly, grouping the variables into theoretical models allowed each distinct concept to be tested, controlling for student body characteristics, using related variables as controls whether or not they were significant. This provided more theoretically coherent models. It was believed that extracting the significant variables from each model and running them in a final general model would have removed them from their theoretical context and controls.

However, creating separate models also has some disadvantages. Variables in separate models cannot control for each other, and some results might be different if variables from other models were included. In addition, while separate models can potentially explain much of the parameter variance, variables in other models might add to the explanatory power.

Therefore, this method of creating separate models could also be seen as a good first step towards creating a general model. A recommended second step would be to choose important (not necessarily significant) variables from each model to combine into a general model. This model would test the association of those variables with controls from other models. Then the proportion of variance explained would express the explanatory power of all the selected variables together, instead of measuring their explanatory power separately in each model.

Choosing the final models and variables

Overview

This section discusses the final HLM models that were chosen for this study, how they differed from the study on the NAEP 1986 data,[42] the reasons why the variables were included, and which variables were conditioning variables. The following section contains the details of variable creation and data preparation.

HLM models consist of student-level models that predict outcomes within schools and school-level models that predict the intercepts and coefficients from the student-level models between schools. The purpose of this study was to develop student-level models to predict two student-level mathematics achievement measures—the overall composite math score and the geometry subscale score—for students in grades 4, 8, and 12, and to develop school-level models that would predict the results of these models. Separate HLM analyses were run for each achievement measure (the math score and the geometry score) within each grade level (4, 8, and 12), producing six separate analyses.

For each analysis, this study used student-level models to predict achievement as a function of gender, race-ethnicity, socioeconomic status (SES),[43] and coursetaking within schools, and used school-level models to predict the intercept, or average achievement, and the gender and race-ethnicity coefficients between schools. While the gender and race-ethnicity coefficients were predicted by school-level models, SES and coursetaking were included in the student-level equations as control variables.

These student-level models differed from the earlier study using 1986 data in the following ways. In the present study, course-taking variables were a new addition to the student-level equations in grades 8 and 12. Also, the parameter variance for the SES and course-taking variables was set to 0, or fixed, and these control variables were not modeled. In the earlier study, SES had been left as a random and modeled variable.

Six school-level models were tested in each grade. These models included the student body characteristics model, the school resources model, the classroom instructional methods model, and three school climate models—math attitudes, student behavior and safety, and academic expectations. The student body characteristics model used exactly the same variables as the earlier study, although school-level SES was measured differently. The school resources model combined variables from the earlier fiscal and physical school characteristics model with variables from the earlier school program structure model. The classroom instructional methods model, the school climate (math attitudes) model, school climate (student behavior and safety) model, and the school climate (academic expectations) model were completely new. Table 2.1 lists the variables that were used in each model.

Only one of the school-level models—student body characteristics—replicated a model tested in the earlier study, and one model—school resources—used variables from

[42]C. Arnold, P. Kaufman, and D. Sedlacek, *School Effects on Educational Achievement in Mathematics and Science: 1985-86* (Washington, DC: U.S. Department of Education, National Center for Educational Statistics, 1992).

[43]SES was created from several variables in the NAEP dataset, some of which were problematic. For more information on this variable, see the discussion of SES in the next section on student-level models.

several of the earlier models. However, most of the models differed from those used in the 1986 study. In addition, more of these models contained variables used for "conditioning" the plausible values of the achievement scores.

Table 2.1.—Variables used in the HLM models

Student-level variables

Dependent Variables

Overall mathematics assessment score
Geometry subscale assessment score

Predictor Variables

All grades
Gender
Race-ethnicity (African-American, Hispanic, Native American, or other versus white or Asian-American)
Socioeconomic status (Constructed from several NAEP variables)

Grade 8 only
Student is currently taking algebra (yes/no)

Grade 12 only
Number of years student has taken geometry
Number of years student has taken calculus

School-level variables, by model

Student body characteristics

Percentage of student body that is African-American
Percentage of student body that is Hispanic
Average socioeconomic level of study sample (aggregated from student-level SES measure)

School resources

Number of students in school
Student/teacher rauo
District instructional funds per student
Microcomputers per student (grade 4)
Percentage of students using computers as part of math instruction (grade 4)

Classroom instructional methods

In math class, in this grade, how often students:
- Work in small groups
- Work with objects (Grade 4: blocks, rulers, shapes; Grade 8: blocks, rulers, solids; Grade 12: rulers, compasses, and protractors)
- Do problems on worksheets
- Do problems from textbook
- Take math tests
- Use calculator
- Use computer
- Write math proofs (grade 12)
- Formulate own problems (grade 12)

Table 2.1.—Variables used in the HLM models—Continued

School climate: Math attitudes

Students feel math is useful
Students enjoy and feel competent in math
Students disagree that math is more for boys (yes/no)

School climate: Student behavior and safety

Index of problems in the school (grade 4)
Percentage of students enrolled from beginning to end of year (grade 4)
Average absenteeism in grade last month (grades 8, 12)
Students feel classes are often disrupted (grades 8, 12)
Students feel unsafe at school (grades 8, 12)

School climate: Academic expectations

Amount of instruction 4th graders receive in math per week
Math is identified as a special priority (yes/no) (grade 4)
Mean composite math score of grade/age sample in school (not used to predict intercept)
Percentage of 8th grade students who are taking algebra
Percentage of 12th grade students in academic/college prep high school program
Mean number of years 12th grade students have taken calculus

SOURCE. U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, *1990 Secondary-use Data Files: Data File Layouts and Codebooks for Age 9/Grade 4, Age 13/Grade 8, and Age 17/Grade 12.*

Student-level models within schools

Dependent variables: Math composite and geometry subscale

Two dependent variables were used in each of the three grades—the student composite score on the total math assessment and the student score on the geometry subscale of the math assessment. Although the geometry score was one of the subscales in the math composite score, this score was examined separately because it represented a higher order math subject in which females and minorities often lag behind males and whites. In addition, geometry scores were available in all three grades, unlike algebra and other higher level subscales.

Independent variables of interest: gender and race-ethnicity status

Within schools, the major variables of interest as predictors of math and geometry achievement for each grade were gender and race-ethnicity. The purpose of the gender variable was to examine the differences in achievement between females and males. The race-ethnicity variable compared the achievement of African-Americans, Hispanics, and Native Americans to that of whites and Asian-Americans. Whites and Asian-Americans were grouped together because the average NAEP scores of these groups were similar and the average scores of the other groups were all much below whites and Asian-Americans.[44]

[44]See J. A. Dossey et al., *The Mathematics Report Card: Are We Measuring Up? Trends and Achievement Based on the 1986 National Assessment* (Princeton, NJ: Educational Testing Service, 1988), and I. V. S.

In addition, Asian-Americans often averaged higher scores than whites, and the purpose of the race-ethnicity variable was to examine the school effects on the achievement gap between the whites and groups who averaged lower scores than whites. The white/Asian-American group was used as the reference group.

These variables measured the differences in achievement by gender and race-ethnicity within schools, and these differences then became dependent variables in between-school models that identified the school characteristics associated with those differences. The gender and race-ethnicity variables were both conditioning variables.

Independent variables as control variables: SES

Also included in the within-school equations in grades 4, 8, and 12 was a measure of student socioeconomic status (SES). This variable was created as an indicator of student disadvantage, and it was included so the association of school characteristics with achievement and with differences by gender and race-ethnicity could be examined holding disadvantage constant. For this reason, it was used as a control variable. While it was associated with student achievement itself, this association was not examined in this study. In the earlier study, few variables were found to predict SES, and this area would be a good one for future HLM research. However, in this study, SES was used as a control variable and was not modeled. In addition, its parameter variance was set to 0, or "fixed," because it tends to be constant across schools. This also provided more degrees of freedom for the gender and race-ethnicity coefficients. The components of SES were also conditioning variables.

SES was created by combining several student-reported variables—mother's and father's education levels and four items about the number of reading items in the home. Because these items were all student-reported, they were not completely reliable, especially the fourth graders' knowledge of parental educational levels. For instance, the parent's educational levels were missing for both parents for 38 percent of fourth grade students, compared to 10 percent of eighth graders and 3 percent of twelfth graders. In contrast, missing values on the four reading items ranged from only 1 to 8 percent for fourth graders, 3 to 5 percent for eighth graders, and 1 to 3 percent for twelfth graders. If any of the six items were missing, the SES value was based on any non-missing items, so SES was missing for a student only if all items were missing for that student. Therefore, the components of SES varied by student. Given this unreliabiliy and variation, it was surprising how consistently significant this variable was in each grade.

Independent variables as control variables: course-taking variables

In grades 8 and 12, course-taking variables were also included in the within-school equations as control variables. Differences in math achievement, especially by gender and race-ethnicity, often result from different course-taking patterns. In addition, the higher the level of math courses taken, the better a student can be expected to score on a math or higher level math test. Therefore, the course-taking measures were used as control variables so that the effects of gender and race-ethnicity could be examined above and beyond these effects.

Mullis and L. B. Jenkins, *The Science Report Card: Elements of Risk and Recovery. Trends and Achievement Based on the 1986 National Assessment* (Princeton, NJ: Educational Testing Service, 1988).

In grade 8, the course-taking measure consisted of whether or not students were taking algebra in the 8th grade, as opposed to taking other 8th grade math course options such as 8th grade math or pre-algebra. Taking algebra controlled for the highest level of current math instruction to which the students had been exposed. In grade 12, different course-taking measures were used to predict geometry and math achievement. In the equation predicting geometry achievement, the number of years of geometry taken by the 12th grade students was used as the control variable because it would control for the amount of geometry to which the students had been exposed. In the equation predicting math achievement, the number of years of calculus taken by the 12th grade students was used as the control variable. Calculus was used because it was one of the highest math-level subjects, so it would control for students who had taken very high levels of math. All of the course-taking variables were conditioning variables.

Since course-taking patterns are associated with differences in achievement by gender and race-ethnicity in the literature, the course-taking coefficients in the within-school equations could be used as dependent variables in equations that would identify school factors associated with these patterns. While this exploration was not the major focus of this study, it would be a good area for future research. However, in this study, these variables were used as control variables only, and they were not modeled. In addition, the parameter variance of these variables was set to 0, or "fixed," in order to provide more degrees of freedom for the gender and race-ethnicity coefficients. However, an exploratory analysis was performed without including coursetaking in the within-school model to determine how much of the gender gap coursetaking explained. The results of this analysis are reported in chapter six.

Between-school models

Student body characteristics

The student body characteristics model was the first school-level model in this analysis to predict average achievement, the gender gap, and the race-ethnicity gap. The student body characteristics used were the racial-ethnic distribution and socioeconomic level of the students. The racial-ethnic distribution measures included the percentage of African-American students and the percentage of Hispanic students in the entire school. The SES measure created at the student level was averaged by school and used as an indicator of the socioeconomic level of the school. This school-level SES variable differed from the 1986 school-level "disadvantaged index." The components of the disadvantaged index used in that study were not available in the 1990 NAEP.

As in the earlier study, these student body characteristics were used both as predictors and as control variables in the between-school models. These variables provide demographic portraits of schools that are often used in distinguishing schools. Therefore, they were first used as predictors to determine whether schools that vary on these aspects also vary in average achievement, gender differences, and race-ethnicity differences. Second, they were used as control variables to ensure that the associations of other school characteristics with average achievement, gender differences, and race-ethnicity differences held true for schools with each type of student body distribution. All subsequent models included these student body characteristics as control variables.

The student body characteristics variables were all conditioning variables. The race-ethnicity variables were conditioning variables at the school level, and the components of the SES variable and the course-taking variables were conditioning variables at the student level, aggregated to the school level.

School resources: Fiscal, physical, staffing, and students

The school resources model tested whether any of the fiscal, physical, staffing, or student resources of schools were associated with average achievement within schools and with gender and race-ethnicity differences within schools. The school resources tested in all three grades were the number of students in the school, the student/teacher ratio, and the amount of district instructional funds per student. In addition, in grade 4, the number of computers per student in the school and the percentage of students in the school using computers as part of math instruction were also tested in an additional school resources model. Except for the percentage of students using computers, all of these variables were tested in the earlier study. However, none were conditioning variables in either study.

The percentage of students using computers was seen as a resource variable because it was an indicator of the availability of computers in the school, i.e. the *physical presence* of computers, while the amount of time spent using computers was seen as an instructional methods variable (next model) because it was an indicator of how time was used for math instruction, i.e., the *time spent using* computers.

Classroom instructional methods

The classroom instructional methods model was based on the amount of time sampled students in each grade reported spending in math class in a number of instructional situations, aggregated to the school level. These instructional situations were working in small groups, working with rulers, blocks, and shapes (grade 4); or rulers, blocks, and solids (grade 8); or rulers, compass, and protractor (grade 12); doing problems on worksheets, doing problems from a textbook, taking math tests, using a calculator, using a computer; and in grade 12, writing math proofs and formulating their own math problems. All of these variables were introduced after the earlier study, and all were conditioning variables at the student level, aggregated to the school level. This model tested whether the average time sampled students in that grade spent in these instructional situations was associated with the level of average achievement and the size of the gender and race-ethnicity gaps.

School climate: Math attitudes

The school climate attitudes model tested how the attitudes students in the sample grade held about math, about themselves and math, and about females and math were associated with average achievement and the size of the gender and race-ethnicity gaps. The attitudes about math were measured by the school average of whether students felt math was used in jobs and useful for solving everyday problems. The attitudes about math and themselves were represented by the school average of whether students liked math and felt they were good at it. Positive attitudes about females and math were expressed by the percentage of students who disagreed that math was only for boys. These attitudes were all measured and used as conditioning variables at the student level, aggregated to the school level. All of these variables were new to this study.

School climate: Student behavior and safety

The school climate student behavior and safety model used the math attitudes variables as control variables, and tested the association of behavior and safety issues with average achievement and the size of the gender and race-ethnicity gaps. Different behavior and safety issues variables were used for grade 4 and for grades 8 and 12. In grade 4, behavior and safety issues were represented by two school variables—an index of problems in the school as reported by the principal, and the percentage of students in the school who were enrolled for the entire school year, which was an indicator of the transient nature of the student body. These variables were measured at the school level and were not conditioning variables. In grades 8 and 12, behavior and safety issues were measured by the average absenteeism of the students, and the percentage of students who do not feel safe at school and the percentage who feel that students often disrupt classes. These variables were measured at the student level in the sample grade and were all conditioning variables, aggregated to the school level.

School climate: Academic expectations

The school climate academic pressures and expectations model used the math attitudes variables as control variables, and tested the association of academic expectations with average achievement and the size of the gender and race-ethnicity gaps. Different measures of academic expectations were used for each grade. In grade 4, academic expectations were measured by the amount of math instruction per week that 4th graders receive in math, and whether math achievement was identified as a special priority in the school. In addition, the school's mean math composite score for grade 4 was used to predict the gender and race-ethnicity gaps. These three variables were measured at the school level, but only the school's math mean was a conditioning variable. In grade 8, the only academic expectations variable was the percentage of 8th graders taking algebra, which is the student-level variable used in the within-school equation aggregated to the school level. This variable was a conditioning variable at the student level. In addition, as in grade 4, the school's mean math composite score for grade 8 was used to predict the gender and race-ethnicity gaps. In grade 12, two variables measured academic expectations—the percentage of students enrolled in the academic/college prep high school program, and the average number of years that students have taken calculus. These variables were student-level variables that were conditioning variables, aggregated to the school level. In addition, as in grades 4 and 8, the school's mean math composite score for grade 12 was used to predict the gender and race-ethnicity gaps.

Normally, average achievement could not be used as an independent variable in any part of an HLM model if the dependent variable was achievement—the equation could not be estimated with such a self-referencing variable included. However, this measure of average achievement was the NAEP-calculated mean of the NAEP grade/age sample. Since that mean was substantially different than the mean that would be obtained from the study sample of the target grade students, the NAEP mean could be included as an alternative measure of academic math ability in the school without jeopordizing the equation. The purpose of including it in the gender and race-ethnicity equations was to test the association of overall math abilities in the school with the gender and race-ethnicity gaps found in the particular target grade.

Checking for collinearity

All of the student-level and school-level variables were checked for collinearity, whether or not they were in the same models, and if they were highly correlated, they were not included in any model. None of the variables in the student-level models were correlated. At the school level, most variables that were highly correlated were rejected for the models. For instance, one classroom instructional method, working autonomously, was highly correlated with all the other methods in every grade, so it was not included. In the grade 12 math attitudes and academic expectations model, the percentages who were on the academic track and taking calculus were chosen because they were the least correlated of any similar variables (r=.37).

However, some variables that were moderately correlated (over r=.50 and significant) were included in the same model. In grade four, math attitudes about math being useful and enjoying math were each correlated with positive attitudes towards females and math (r=.70). In grade 8, the aggregated SES composite variable was moderately correlated with percentage of Hispanics in a school (r=-.50), the school math mean was correlated with SES (r=.70), and somewhat correlated with percent African-American in a school (r=-.55). As in grade four, grade 8 math attitudes about math being useful and enjoying math were each correlated with positive attitudes towards females and math (r=.76 and r=.91, respectively).

In grade 12, the school math mean was somewhat correlated with percent African-American in a school (r=-.57). This school math mean was also somewhat correlated with aggregated SES (r=.72), with aggregated not feeling safe (r=-.52), and with percentage on academic track (r=-.59). Aggregated time spent using textbooks was moderately correlated with taking tests (r=.57) and using calculators (r=.61). As in grades 4 and 8, grade 12 math attitudes about math being useful and enjoying math were each correlated with positive attitudes towards females and math (r=.82 and r=.91, respectively). In addition, attitudes about math being useful and those about enjoying math were somewhat correlated (r=.52). While correlations between .50 and .79 should not affect the models, the correlations above .80 in the math attitudes variables may have made those equations in grades 8 and 12 somewhat unstable.

Other 1990 NAEP variables considered for models and not used

Many variables in the 1990 NAEP were expected to be used in the models, but they were not included because either they had too many missing values, their interpretation in an HLM model would be unclear, or their meaning was unclear.

In grades 4 and 8, almost all of the teacher variables had missing values for more than 10 percent of the cases, and most were missing 15 or more percent. These teacher variables in grades 4 and 8 (no teacher variables were collected for grade 12) would have provided valuable information on school and classroom atmosphere and instructional methods. Fortunately, there were student-level measures of attitudes and instructional methods that could be aggregated to the school level to serve as school-level proxies of school climate and instructional methods. However, the student-reported measures of time spend on various instructional methods were probably not as reliable as teacher reports would have been.

In grades 8 and 12, many school and principal variables, including many of the school resources variables used in the 1985–86 study, and many principal characteristics that would have been especially pertinent for studying gender and race-ethnicity achievement differences, had missing values for 15 percent or more of the cases. Since the grade 4 dataset had fewer school and principal variables with missing values, a detailed HLM model of school resources was developed for grade 4 with variables that could be used in grade 4 but not in grades 8 or 12. In addition, grade 4 principal variables were tested in a separate model; however, they were not reported because they were not significant and could not be compared with grades 8 and 12. Some variables, such as the number of math specialists and aides, were missing too many cases to be used in any grade.

Some variables were not used because their HLM interpretation would have been unclear. For instance, if watching TV had been correlated with average achievement, it is not known whether this variable was measuring leisure time, time that could have been spent studying, or something else. In addition, the meaning of some variables was uncertain. The rural/suburban/urban distinctions were unclear enough not be able to trust using this variable. Since the community economic variables such as percent in occupational groups and the *Orshansky* percentile did not provide the actual income levels of the students in particular schools, they were not useful as measures of SES in the students' particular commuity. Therefore, the school average of the constructed student SES variable was used instead.

Variable construction

The variables used in this analysis are listed in table 2.2. Field names from the appropriate NAEP data file are provided in table 2.1 for those variables used directly from the files. "Composite" in the field name column indicates that the variable was created for this analysis from several other variables. "Dummy" in the field name column indicates that the variable was transformed into one or more dummy variables. Composite and dummy variables are described in table 2.3.

Table 2.2.—NAEP variables used in HLM models

Field name	Variable label
	Student-level variables
Dependent Variables	
MRPCMP1	Overall math assessment score: NAEP plausible value 1
MRPCMP2	Overall math assessment score: NAEP plausible value 2
MRPCMP3	Overall math assessment score: NAEP plausible value 3
MRPCMP4	Overall math assessment score: NAEP plausible value 4
MRPCMP5	Overall math assessment score: NAEP plausible value 5
MRPSCC1	Geometry subscale score: NAEP plausible value 1
MRPSCC2	Geometry subscale score: NAEP plausible value 2
MRPSCC3	Geometry subscale score: NAEP plausible value 3
MRPSCC4	Geometry subscale score: NAEP plausible value 4
MRPSCC5	Geometry subscale score: NAEP plausible value 5
Predictor Variables	
Dummy	Gender
Dummy	Race-ethnicity
Composite	Socioeconomic status
Dummy	Student is currently taking algebra (grade 8 only)
M811005B (order reversed)	No. of years student has taken geometry (gr. 12 only)
M811011B (order reversed)	No. of years student has taken calculus (gr. 12 only)
	School-level variables
Student body characteristics	
PCTBLKQ	Percentage of student body that is African-American
PCTHSPQ	Percentage of student body that is Hispanic
Composite (aggregated)	Average socioeconomic level of grade sample
School resources	
SNSTUDA	Number of students in school
SNSTUDA/SNTCHA	Student/teacher ratio
SIDP	District instructional funds per student
C029601/SNSTUDA	Microcomputers per student (grade 4)
C028001	Percentage of students using computers in math (grade 4)

Table 2.2.—NAEP variables used in HLM models—Continued

Field name	Variable label
Classroom instructional methods	
	In math class in this grade, how often students:
M810103B (order reversed, aggregated)	Work in small groups
M810112B (order reversed, aggregated)	Work with objects
M810102B (order reversed, aggregated)	Do problems on worksheets
M810101B (order reversed, aggregated)	Do problems from textbook
M810107B (order reversed, aggregated)	Take math tests
M810105B (order reversed, aggregated)	Use calculator
M810106B (order reversed, aggregated)	Use computer
M810109B (order reversed, aggregated)	Write math proofs (grade 12)
M810110B (order reversed, aggregated)	Formulate own problems (grade 12)
School climate: Math attitudes	
Composite (aggregated)	Students feel math is useful
Composite (aggregated)	Students enjoy and feel competent in math
Dummy (aggregated)	Percentage who disagree that math is more for boys
School climate: Student behavior and safety	
Composite	Index of problems in the school (grade 4)
C028301	Percentage of students enrolled all year (grade 4)
S004001A (aggregated)	Average absenteeism in grade last month (grades 8, 12)
B007003A (aggregated)	Students feel classes are often disrupted (grades 8, 12)
B007002A (aggregated)	Students feel unsafe at school (grades 8, 12)
School climate: Academic expectations	
C030102	Amount of math instruction per week (grade 4)
C029203	Math is identified as a special priority (yes/no) (grade 4)
SMEANM	Mean composite math score of grade sample
Dummy (aggregated)	Percentage taking algebra (grade 8)
Dummy (aggregated)	Percentage in academic/college prep program (grade 12)
M811011B (order reversed, aggregated)	Mean years students have taken calculus (grade 12)

NOTE: Order reversed: the order of the values in the original NAEP variable was reversed. Aggregated: student-level variable(s) were averaged across students within schools and aggregated to the school level.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, *1990 Secondary-use Data Files: Data File Layouts and Codebooks for Age 9/Grade 4, Age 13/Grade 8, and Age 17/Grade 12.*

The specific variables included in each composite and dummy variable are shown in table 2.3. If the component variables were standardized for purposes of averaging, this is indicated under the variable name. If the variables were student-level variables that have been aggregated to the school level, this is indicated under the field name. A discussion of how these variables were created and constructed follows the table.

Table 2.3.—Composite and dummy variables used in HLM models

Variable name	Field name	Variable label
Student-level variables		
Gender	DSEX	Females=1, Males=0
Race-ethnicity	DRACE	Minority=1 (DRACE=African-American, Hispanic, or American Indian) Non-minority=0 (DRACE=white or Asian)
Socioeconomic status (averaged)	B003501A (standardized) B003601A (standardized) B000901A (standardized) B000903A (standardized) B000904A (standardized) B000905A (standardized)	Mother's education Father's education Does family get newspaper regularly Is there an encyclopedia in home Are there more than 25 books in home Does family get magazines regularly
8th grade student is currently taking algebra	M810501B	1=taking algebra in 8th grade 0=taking no math, 8th grade math, algebra, pre-algebra, or other in 8th grade
School-level variables		
Average socioeconomic level of grade sample (averaged & aggregated)	B003501A (standardized) B003601A (standardized) B000901A to B000905A (standardized)	Student-level SES measure as defined above, aggregated to school level
Students feel math is useful (averaged and aggregated)	M811102B (grade 4) M810702B (grades 8, 12)	Student feels that all people use math in jobs (order reversed)
	M811105B (grade 4) M810705B (grades 8, 12)	Student feels that math is useful for everyday problems (order reversed)
Students enjoy and feel competent in math (averaged & aggregated)	M811101B (grade 4) M810701B (grades 8, 12)	Student likes math (order reversed)
	M811103B (grade 4) M810703B (grades 8, 12)	Student feels he or she is good in math (order reversed)
Percentage of students who disagree that math is more for boys (aggregated)	M811104B (grade 4) M810704B (grades 8, 12)	1=Disagree (grade 4); or disagree/strongly disagree (grades 8, 12) that math more for boys 0=agree or undecided (grade 4); or strongly agree, agree, or undecided (grades 8, 12) that math more for boys

Table 2.3.—Composite and dummy variables used in HLM models—Continued

Variable name	Field name	Variable label
Index of problems in the school (averaged)	C028201	Student tardiness (order reversed)
	C028202	Student absenteeism (order reversed)
	C028203	Student cutting of classes (order reversed)
	C028204	Physical conflicts among students (order reversed)
	C028205	Robbery or theft (order reversed)
	C028206	Vandalism of school property (order reversed)
	C028207	Student use of alcohol (order reversed)
	C028208	Student use of illicit drugs (order reversed)
	C028209	Student possession of weapons (order reversed)
	C028210	Physical abuse of teachers (order reversed)
	C028211	Verbal abuse of teachers (order reversed)
Percentage of 8th graders taking algebra (grade 8) (aggregated)	M810501B	1=taking algebra in 8th grade 0=taking no math, 8th grade math, pre-algebra, or other in 8th grade
Percentage of 12th graders in academic/college prep program (grade 12) (aggregated)	B005001A	1=High school program best described as academic/college prep 0=High school program best described as general or vocational/technical

NOTE: Composite variables are assigned missing values only if all component variables are missing. Order reversed: the order of the values in the original NAEP variable was reversed. Aggregated: student-level variable(s) were averaged across students within schools and aggregated to the school level. Averaged: Components of composite variables were averaged within each student or school.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, *1990 Secondary-use Data Files: Data File Layouts and Codebooks for Age 9/Grade 4, Age 13/Grade 8, and Age 17/Grade 12.*

Creation of Dummy Variables

Five dummy variables were created from the NAEP variables. The derived NAEP variable for gender was used to create the gender variable by changing the codes to make males the reference group. The derived NAEP variable for race-ethnicity was changed into a dummy variable by designating African-Americans, Hispanics, and American Indians as one group, and whites and Asian-Americans as another group.[45]

The student-level NAEP variable for attitude about gender and math, M811104B (grade 4) and M810704B (grades 8 and 12), stating that math is more for boys, was converted to a dummy variable before being aggregated to the school level. The categories of "disagree" (grade 4) or "disagree" and "strongly disagree" (grades 8 and 12) were put into one group that designated positive feelings about females and math; the categories of "agree," "undecided," and "strongly agree" were put into the reference group. Since this dummy variable was aggregated to the school level, it designated the percentage of students in that grade in each school who had positive feelings about females and math.

[45]There are only five categories of race-ethnicity in NAEP.

The NAEP student-level variable for whether 8th graders who were taking algebra in 8th grade, M810501B, was already a dichotomous variable with a value of 1 if the students were taking algebra. The reference group value was changed from 2 to 0. This variable was then aggregated to the school level, designating the percentage of 8th graders who were taking algebra in each school. Similarly, the student-level NAEP variable for whether 12th graders were in an academic/college prep track or not was already dichotomous, and the reference group was simply changed to 0. When this variable was aggregated to the school level, it represented the percentage of 12th graders in the academic track in each school.

Construction of continuous variables

The continuous composite variables were constructed in the following manner. The student-level SES variable was created by combining several student-reported responses—mother's and father's education levels and several items about the number of reading items in the home. If these or any other items were missing, the SES value was based on any non-missing items. SES was missing for a student only if all items were missing for that student. Each non-missing component of SES was standardized, and the mean standardized score of these components became the SES value. Therefore, the components of SES varied by student. For the school-level SES variable, the student-level SES values were averaged within each school.

For the math attitude variables, whether math is useful and whether the students like and enjoy math, the two variables were combined. These variables were created first by reversing the order of the scores, and second by shifting the 5-point scale so agree was above 0 and disagree was below 0. Then the average of these two variables was obtained for each student. These scores were averaged for each school to obtain the school-level measures of this attitude (which were really the grade-level measures since they just included the attitudes of the students in the sample grade).

In 4th grade, the index of problems in the school was a NAEP-created school-level variable that averaged the score of each component school-level variable. For this study, the order of the index was reversed so that schools with more problems had higher scores.

Table 2.4 lists the ranges and the unstandardized means and standard deviations of all the school-level variables used in this analysis. In addition, table 2.4 shows the across-school means and standard deviations of the within-school variables of the math and geometry plausible values and gender, race-ethnicity, SES, and coursetaking.

Table 2.4.—Unstandardized means and standard deviations for student-level and school-level independent variables, by grade: 1990

Variable (range)	Grade 4 Mean	Grade 4 (s.d.)	Grade 8 Mean	Grade 8 (s.d.)	Grade 12 Mean	Grade 12 (s.d.)
Student-level variables						
Dependent variables						
Math1 (98-394)	214.27	(29.36)	263.51	(32.61)	294.76	(33.93)
Math2 (76-388)	214.65	(29.19)	263.58	(32.70)	294.58	(33.55)
Math3 (116-390)	214.46	(29.36)	263.81	(32.46)	294.82	(33.69)
Math4 (111-395)	214.75	(29.32)	263.69	(32.73)	294.76	(33.81)
Math5 (112-391)	214.45	(29.47)	263.74	(32.41)	294.66	(33.76)
Geometry1 (100-412)	215.88	(32.37)	260.60	(35.67)	295.33	(42.01)
Geometry2 (92-432)	216.21	(32.20)	260.90	(35.11)	295.18	(41.47)
Geometry3 (100-423)	215.85	(32.64)	261.06	(35.21)	295.72	(41.18)
Geometry4 (103-439)	216.03	(32.13)	260.85	(35.25)	295.70	(41.40)
Geometry5 (101-420)	216.29	(32.64)	260.82	(34.64)	295.70	(41.18)
Predictor variables						
Gender–if female (0/1)	0.48	(0.50)	0.50	(0.50)	0.52	(0.50)
Race-ethnicity–if African-American, Hispanic, or Native American (0/1)	0.29	(0.45)	0.27	(0.44)	0.23	(0.42)
SES level	-0.02	(0.58)	-0.02	(0.61)	-0.01	(0.58)
Taking algebra in grade 8 (0/1)			0.14	(0.35)		
Number of years of geometry by grade 12 (0-3)					1.48	(0.92)
Number of years of calculus by grade 12 (0-3)					0.21	(0.62)
Number of students	5,080		5,198		4,953	
School-Level variables						
Student body characteristics						
Percent African-American (0-100)	13.03	(20.06)	11.74	(19.38)	10.13	(17.22)
Percent Hispanic (0-100)	7.52	(13.63)	5.77	(11.96)	5.12	(10.84)
SES level	0.04	(0.25)	-0.05	(0.25)	-0.04	(0.22)
School resources						
Number of students	472.36	(223.86)	542.76	(314.62)	850.36	(629.46)
Student/teacher ratio	19.60	(6.62)	16.43	(4.63)	17.05	(6.00)
Instructional funds/student (1-9)	6.05	(1.66)	5.97	(1.71)	5.99	(1.60)
Microcomputers/student	0.06	(0.05)				
Percent students using computer (0-100)	75.12	(30.73)				

Table 2.4.—Unstandardized means and standard deviations for student-level and school-level independent variables, by grade: 1990—Continued

Variable (range)	Grade 4 Mean	Grade 4 (s.d.)	Grade 8 Mean	Grade 8 (s.d.)	Grade 12 Mean	Grade 12 (s.d.)
Classroom instructional methods						
In math class, how often:						
Work in small groups (1-5)	2.27	(0.66)	2.10	(0.79)	2.20	(0.50)
Work with objects (1-5)	2.49	(0.60)	2.15	(0.59)	2.20	(0.39)
Do problems on worksheets (1-5)	3.65	(0.60)	3.03	(0.74)	2.55	(0.51)
Do problems from textbook (1-5)	4.12	(0.55)	4.52	(0.50)	3.86	(0.63)
Take math tests (1-5)	2.73	(0.42)	2.84	(0.35)	2.47	(0.35)
Use calculator (1-5)	1.72	(0.51)	2.41	(1.07)	3.38	(0.69)
Use computer (1-5)	2.19	(0.60)	1.61	(0.48)	1.85	(0.45)
Write math proofs (1-5)					1.97	(0.35)
Formulate own problems (1-5)					1.54	(0.27)
School climate: Math attitudes						
Math is useful (-2- +2)	0.51	(0.19)	1.02	(0.22)	0.85	(0.25)
Enjoy & feel competent in math (-2- +2)	0.51	(0.19)	0.49	(0.37)	0.35	(0.34)
Math not more for boys (0-1)	0.82	(0.12)	0.75	(0.09)	0.81	(0.11)
School climate: Student behavior and safety						
Problems in the school (0-3)	0.42	(0.31)				
Percent enrolled full year (0-100)	82.23	(23.18)				
Average days absent in month			1.85	(0.28)	2.20	(0.35)
Students feel classes often disrupted (1-4)			2.91	(0.29)	2.59	(0.28)
Students feel unsafe at school (1-4)			1.85	(0.28)	1.68	(0.28)
School climate: Academic expectations						
Instruction/week in math (1-5)	3.71	(0.86)				
Math special priority (0/1)	0.79	(0.39)				
Mean composite math score	213.90	(15.50)	261.85	(18.32)	292.31	(18.10)
Percent taking algebra in grade 8 (0-1)			0.12	(0.13)		
Percent of grade 12 in academic/ college prep (0-1)					0.52	(0.21)
Average number years of calculus by grade 12 (0-3)					0.18	(0.19)
Number of schools	257		174		186	

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th, 8th, and 12th Grade Students, Restricted-Use Data Base.

Data and variable preparation

The HLM analysis and software program require researchers to make many decisions about data and variable preparation before and during the HLM analysis. Some of these decisions simply affect the ability of the HLM software to handle the data; others affect the interpretation of the results. Some procedures were necessary due to the structure of the NAEP datasets; others would be required in any HLM analysis.

This section discusses the data preparation decisions made in this analysis and the implications of those decisions for interpretation. The major data and variable preparation procedures were weighting, preparing missing values, centering, standardizing, and working with any software limits on the numbers of variables for the models.

Weighting

The data were weighted at the student and school level, and the weights at each level were normalized. Both the multi-stage sampling plan of NAEP and its associated sampling error need to be taken into account when estimating parameters. Normal analysis of NAEP data uses the jackknife weights. However, the multi-level nature of HLM models (students within schools) directly models the NAEP design of sampling students within schools. Therefore, the modeling that is performed in HLM accurately reflects the sampling design.[46] Use of the appropriate weight at each level of the analysis provides accurate population estimates for each variable.[47] HLM can also distinguish between the sampling error and measurement error, so that the contribution of sampling error to the overall error can be identified.

These analyses were weighted using both the student and school weights provided by NAEP to reflect the sampling design and response rates. Because HLM uses the weights from both student and school levels at the same time, new within-school student weights were calculated by dividing out the school weight from the original student weights. Using the original student weights would have resulted in the school weights being counted twice. The weights were normalized so they would provide the same proportionate weighting of each case, but would sum to the unweighted sample size. Using the actual weights would have produced a sample that was inappropriately large for the HLM statistical tests. The current HLM PC version provides an option to normalize the weights within the HLM program if the researcher has not done so already.

Missing values

HLM allows missing values in the within-unit variables (that is, at the student level). There were no missing values in the gender or race-ethnicity variable, but the missing values in the SES variable reduced the within-school cases considerably, sometimes even to the point of eliminating the entire school from the analysis.

---

[46]This is actually only partly true. The primary sampling unit level of analysis is not in the model. However, one would have to jackknife the entire HLM analysis to see how much of the design effect is reflected in the standard error of estimate for HLM.

[47]These estimates are not perfectly accurate because of the students and schools that were dropped from the analysis due to missing data and subsample selection. See the following section.

Since HLM does not allow missing values in the between-unit variables, schools with missing values on these variables were handled in two different ways. First, schools with missing values on a variable were assigned the mean of that variable across all schools. If only one or two schools, or cases, had missing values, then this mean served as the value for that variable for those cases. A second approach was used if there were more than 2 cases but not more than 10 percent of the cases with missing values: a "missing value dummy variable" indicated with a "1" that the case had a missing value on that variable, while all other cases were assigned a "0" on that dummy variable. In other words, the mean of that variable for cases with missing values served as just a place holder. Thus, these missing value dummy variables controlled for missing values in the equation and at the same time included those cases in the equation. However, these dummy variables added to the number of variables in each model. Variables with more than 10 percent missing values were not used in the analysis.

The missing value dummy variables are reported in the Appendix A tables. However, they are not included in the HLM results tables (chapters three through five) because they do not affect the results.

Centering

The student-level variables—gender, race-ethnicity, SES, and coursetaking—were centered, that is, their school means were subtracted from them so their means were all 0. This allowed the *intercept* to be interpreted as the average achievement in each school, not controlling for the other within-school variables because they were centered.

However, the gender coefficient could then be interpreted as the gap between females and males (the "gender gap") in each school, controlling for race-ethnicity, SES, and coursetaking. The race-ethnicity coefficient could be interpreted as the gap between African Americans/Hispanics/Native Americans and whites/Asians (the "race-ethnicity gap") in each school, controlling for gender, SES, and coursetaking. Since dummy variables were used for gender (female=1; male=0) and race-ethnicity (African-Americans, Hispanics, and Native Americans=1; whites and Asian-Americans=0), the mean of these values was the percentage of females or minorities in that school. The difference between the two values on either variable was still 1, with females and minorities having the positive values, and males and whites/Asian-Americans having the negative values.

Since SES already had a 0 mean, the SES coefficient could be interpreted as indicating the extent to which SES was associated with achievement in each school, controlling for gender, race-ethnicity, and coursetaking. The coursetaking coefficient was interpreted as the extent to which math coursetaking was associated with math achievement in each school, controlling for gender, race-ethnicity, and SES.

The main reason for centering was to be able to interpret the intercepts of the within-unit equations in the following way. Since the intercept was the average level of achievement in a school when the three or four predictors were at 0, and since 0 was their mean, the intercept was the level of average achievement in each school at "average" gender, race-ethnicity, SES, and coursetaking. Although there is no real "average" gender or race-ethnicity, this achievement level can be seen as the average achievement before the effects of gender, race-ethnicity, SES, and coursetaking have been taken into account. Since the intercept becomes the dependent variable in the first between-school equation, this equation can be interpreted as predicting the average achievement in each school overall, rather than for some limited group, such as the achievement of white males of average SES with high coursetaking. This provides a baseline, if hypothetical, level of average achievement, which the parameters of gender, race-ethnicity, SES, and coursetaking can then alter.

While centering did not change the value of the Beta coefficients of gender, race-ethnicity, SES, or coursetaking, it did allow a more descriptive interpretation of these coefficients. In the case of the dummy variables, the coefficients still represented the average difference in the number of achievement points between males and females, and between African-Americans/Hispanics/Native Americans and whites/Asian-Americans. If the coefficients were positive, the females and African-Americans/Hispanics/Native Americans were doing that much better than males and whites/Asian-Americans, i.e., the gap was smaller. If the coefficients were negative, females and minorities were doing that much worse, i.e., the gap was larger.

Instead of seeing the coefficients as the values for females or minorities, these same coefficients were interpreted as the "gap" between females and males, or between African-Americans/Hispanics/Native Americans and whites/Asian-Americans, since 0 was not males or white/Asian-Americans, but somewhere between the dichotomous values. These Beta coefficients, or parameters, are the dependent variables in the between-school equations, and will be referred to in the text as the "gender gap" between girls and boys in achievement, or the "race-ethnicity gap" between African-Americans/Hispanics/Native Americans and whites/Asian-Americans in achievement.

In the case of SES, the continuous variable, its value was positive above its mean (0) and negative below its mean, instead of going from 0 to a higher value. A positive SES coefficient would push the SES value away from 0 in either direction, pushing the achievement level in the corresponding direction and creating a larger difference in achievement between students of high or low SES. A negative coefficient would push the SES value toward 0 from either direction, reducing the change in achievement level and creating a smaller difference in achievement between students of high or low SES. The Beta coefficient on SES could thus be interpreted as the "differentiating effect" of SES, and will be referred to in this way in the text.

In the case of coursetaking, the coefficient of the dummy variables of taking 8th grade algebra represented the average difference in achievement points between those who were taking algebra and those who were not. The interpretation of the coefficients of the number of years the 12th grade students had taken geometry and calculus could be interpreted like the SES coefficient, that is, as the differentiation effect of years of calculus or geometry.

An issue in centering in HLM models is whether and how to include the school means of each of the centered within-unit variables in the between-unit equations. It is generally agreed that they should be included, unless the researchers want all the schools to be treated as if they have the same means on these variables, since all of these means have been set to 0.[48] If the school means are to be included, it must be decided whether to include the means for each school from the sample, or to use school means from another source. The most accurate source is recommended.

In the case of the NAEP school datasets, school-level estimates of the percentage of African-American and Hispanic students and variables that indicate the SES of the

[48]For a technical discussion of these and other centering issues see S. W. Raudenbush, "'Centering' Predictors in Multilevel Analysis: Choices and Consequences." *Multilevel Modeling Newsletter* 1 (2) (1989): 10–12; N. T. Longford, "To Center or Not to Center," *Multilevel Modeling Newsletter* 1 (3) (1989): 7; I. Plewis, "Comment on 'Centering' Predictors in Multilevel Analysis, *Multilevel Modeling Newsletter* 1 (3) (1989): 8–10.

community were potential choices for school means. The school-level race-ethnicity estimates were a more accurate estimate of school-level race-ethnicity than simply averaging the within-school race-ethnicity from the small samples of students in the dataset in each grade. However, the school-level estimates of SES were very different from the within-school student level SES variable; thus, aggregating the student SES level of the grade sample to the school level was seen as a more accurate school-level estimate of SES. These variables, the percentage of African-American and Hispanic students in the school and the aggregated SES level of students in that grade, constituted the first HLM model tested in this study, and all of these variables were included in all subsequent models.

There were no single-sex schools since the sample was of public schools only, so the gender mean was assumed to be constant at 51 percent and was not included. However, this illustrates the dilemma of wanting to center a within-unit independent variable in order to make the intercept of the within-unit equation a true average without having a between-unit measure of the mean of that variable. This issue will require more discussion among HLM researchers.

In the case of coursetaking in grades 8 and 12, there were no school-level estimates of coursetaking; therefore, these had to be aggregated from the student-level variables. However, they were only included in one of the models. Since these variables were fixed at the within-school level and not modeled, it was not important to adjust their means.

Standardizing

Whether or not to standardize the independent or dependent variables is the choice of the researcher. In this study, all the school-level variables in this study were standardized, so their values were in standard deviation units from their mean.[49] However, the dependent variables—the achievement scores—were not standardized in this study. Therefore, the Gammas in the between-school equations were interpreted as the change in the dependent parameter, in numbers of points on the achievement score scale, for every standard deviation above the mean of each school-level variable.

To understand standardized variables, start with a regular one level linear regression equation. As in regular regression models, the independent school-level variables with significant coefficients in HLM models are interpreted as predicting, for every unit change in that variable, a change in the dependent variable (in this case the Beta coefficient or intercept) by the amount of the Gamma coefficient. When the between-school variables are standardized to a mean of 0 and a standard deviation of 1, their unit changes are in standard deviation units. The coefficients of the between-school variables, the Gammas, thus predict how much the dependent variable will change for every standard deviation of these between-school variables. This change is predicted for every level of (that is, controlling for the effects of) the other independent variables in the equation.

The choice of standardizing the school-level variables was made in order to allow the size of the Gamma coefficients on these variables to be comparable within each model so the variables with the largest coefficient, or association, with the dependent variable—the within-school intercept or Beta coefficient—could be identified. For policy recommendations, the ability to determine the variables with the largest as well as strongest association with average school achievement or the gender or race-ethnicity gap in

[49]The missing value dummy variables were not standardized.

achievement was believed to be more important than the actual difference in achievement points.

However, the meaning of the size of standardized variables is often difficult to interpret when the school-level variables are not in their original units. The unstandardized means and standard deviations can provided (as they are in Table 2,4), to convert the variables of interests in their original units. However, some researchers and readers prefer independent variables to be unstandardized in the first place. Then, the Gamma coefficients can be interpreted as the change in the dependent parameter for every actual unit of each school-level variable. Although the size of these coefficients cannot be compared between variables, the meaning of the size of the coefficient for each individual variable is clearer because it is expressed in its original units.

The student-level dependent variables—the achievement scores—were not standardized in this study. Standardizing the outcome variable provides meaning to the value of the Gammas for the dependent variable. Whether the change in the independent variable is in standardized or unstandardized units, the value of the Gammas can be interpreted as the number of standardized units of the achievement score, so the meaning of a difference of several points on the achievement score scale can be more easily understood. This was not done in this study because the number of points on the achievement score scale were thought to have meaning due to the anchoring by grade of the NAEP scores. In addition, standardizing the independent variables had already removed the reader one step from the data. Interpreting the meaning of the level-2 HLM Gamma coefficients using standardized school-level variables as predictors for the within-school average achievement and the gap in achievement by gender and race-ethnicity was already hard enough in actual achievement score points.

Limits on number of variables

While the new PC version of HLM does not limit the number of variables allowed in the sufficient statistics files and in each equation, this analysis was restricted by what appeared to be a DOS limit on the length of the format statement for the sufficient statistics file. This format statement length limit allowed only about 20 variables to be included in each sufficient statistics file, and necessitated creating two or more sufficient statistics file for each grade. In this case, it was lucky that separate groups of variables (models) were being estimated. Since the models were tested separately, variables were grouped within sufficient statistics files, and the appropriate file was used for each model. Therefore, this limit did not affect the analysis, other than the inconvenience of creating multiple sufficient statistics files.

C. Estimating HLM models with NAEP plausible values

Estimation theory and practice

NAEP used item response theory (IRT) to estimate proficiency scores in math and science for each sampled student. However, the IRT proficiency scores are biased, so conditional proficiency scores are estimated using student background information. These conditional proficiency scores are latent variables conditional on the student's responses to several cognitive and background items and *are not directly observed.* That is, proficiency scores were predicted from a set of cognitive and background variables (referred to as *conditioning variables*). Because the proficiency scores are not observed, but estimated, there is some amount of uncertainty or variance associated with them. Thus, rather than having a single observed math score, there is a range, or distribution, of *plausible values* for each student's proficiency in math. A sample of these plausible values are chosen through random draws from the conditional distribution of proficiency scores for each student. This sample of plausible values is used by researchers to analyze student proficiency scores. The variance in the distribution of plausible values reflects the errors in measurement, and the method of estimating that measurement error, along with the sampling error associated with choosing a sample of those values, is described below.

In the NAEP dataset, NAEP provides a sample of five plausible values for each student. However, for any analysis of student achievement scores, from simple means to complex HLM models, researchers are interested in obtaining one estimate rather than five estimates for each student or group. Therefore, the NAEP *Data Files User Guide*[50] provides directions about how to conduct analyses with the five plausible values. Although it is tempting to do so, NAEP explicitly warns researchers against simply averaging the five values and conducting analyses on that average. This procedure would severely bias the results. Instead, the correct procedure is to conduct the desired analysis on each of the five plausible values, and then to average the parameter estimates from those five analyses.

Therefore, the parameter estimates from the HLM analyses in this report were based on the average parameter estimates from separate HLM analyses of the five plausible values. That is, for each HLM model, a separate HLM analysis was conducted on each of the five plausible values, and the results from these analyses were averaged. The HLM parameter estimates that were averaged for this report included the Gammas, as well as the following statistics that are explained in the following section—the parameter variances (Tau), the reliabilities, the Chi-square test for the parameter variance being 0, and the probability of that Chi-square value.[51] In addition, the standard error of the averaged Gammas was estimated as described below. The Student's *t* value for the Gammas was calculated by dividing the averaged Gamma by its standard error. The probability of this *t* value was estimated from a standard *t* distribution table.

The standard error of the Gammas consisted of two components—sampling error and measurement error. The following routine provided in the NAEP *Data Files User Guide*[52]

---

[50] A. Rogers et al., *National Assessment of Educational Progress: 1990 Secondary-use Data Files User Guide* (Princeton, NJ: Educational Testing Service, March, 1992).

[51] The software developed to use plausible values in an HLM analysis, "HLMPV," calculates the probability of the Chi-square value in a much more sophisticated way, using the equation used in the regular HLM software.

[52] A. Rogers et al., *National Assessment of Educational Progress: 1990 Secondary-use Data Files User Guide* (Princeton, NJ: Educational Testing Service, March, 1992).

was used to approximate the component of error variance in the analysis due to the error in measurement and to add it to the sampling error:

Let $\hat{\theta}_m$ represent the m^{th} plausible value, where m=1 to M sets of plausible values (in our case M=5). Let $\hat{t}_m$ represent the parameter estimate based on the m^{th} plausible value. Let U_m represent the variance of $\hat{t}_m$, or the sampling error.

- Five HLM runs were conducted based on each plausible value $\hat{\theta}_m$. The parameter estimates from these runs were averaged:

$$t^* = \frac{\sum_{m=1}^{M} \hat{t}_m}{M}$$

- The variance of the parameters from these runs were averaged:

$$U^* = \frac{\sum_{m=1}^{M} U_m}{M}$$

- The variance of the M estimates $\hat{t}_m$ was estimated:

$$B_m = \frac{\sum_{m=1}^{M} (\hat{t}_m - t^*)^2}{(M-1)}$$

- The final estimate of the variance of the parameter estimate is the sum of the two components:

$$V = U^* + (1 + M^{-1})\, B_m$$

The square root of this variance is the standard error of the Gamma, and it is used in a standard Student's t formula to evaluate the statistical significance of each Gamma.

As evident from the preceding formula, the use of plausible values in an HLM analysis usually increases the standard errors of the Gamma coefficients, making it harder to identify significant school-level correlates.[53]

---

[53] P. Kaufman, C. Arnold, and M. Wilson, *Using Plausible Values in Hierarchical Linear Models* (Washington, DC: U.S. Department of Education, National Center for Education Statistics, 1991).

Software for HLM estimates with NAEP data

The HLM analyses in this study were produced using HLM/2L, a two-level HLM microcomputer program developed by Anthony Bryk, Stephen Raudenbush, and Richard Congden.[54] However, due to the five plausible values rather than one proficiency score that NAEP provides for each student, the task of producing final estimates for the HLM models involved more than running this program. Usually in NAEP, parameter estimates are produced by estimating each parameter for each of the five plausible values, and by averaging the five estimates. Then, the standard error for this average estimate is calculated with a formula provided in the NAEP technical manual.

In order to follow this procedure with HLM models, for every model, five HLM analyses had to produced, using each of the five plausible values. Then a way had to be found to average the results from those five analyses and produce the correct standard errors. Besides hand calculating the averaged estimates, there are two computerized methods to accomplish this. The first method is to transfer the results to spreadsheet software, and average the results there. The HLM results in this report were produced with spreadsheets. The second method is to use software that has been developed to automatically take the results from the HLM software and produce correctly averaged results. These two methods are described in more detail below.

Averaging with spreadsheets

In an HLM analysis with NAEP data, the parameter estimates are based on the average parameter estimates from separate HLM analyses of the five plausible values. That is, a separate HLM analysis is conducted on each of the five plausible values. The results from the five analyses are averaged, and the standard errors are calculated as outlined above.

To accomplish these calculations with a spreadsheet program, HLM estimates are first produced for each plausible value. Then, all of the final estimates in the HLM output files are extracted and copied into a spreadsheet program. Then, in this program, the estimates are either averaged, or, in the case of the standard errors and t values, calculated with specified formulas. These procedures are straightforward and could even be performed with a hand calculator if just one or two HLM models were produced. However, the more models tested, the more helpful a spreadsheet program becomes. However, even with a spreadsheet these procedures can be tedious and time-consuming when performed on many models, grades, and dependent variables. For this reason, the following software was developed.

New software for the use of plausible values with HLM

HLM2PV (PV for plausible values) is software that integrates HLM/2L, the HLM software, with the procedure of taking NAEP plausible values into account in the HLM estimates. It was developed by the author and a programmer during this study, and the

---

[54] A. S. Bryk, S. W. Raudenbush, M. Seltzer, and R. Congdon, *An Introduction to HLM: Computer Program User's Guide*, 2nd ed. (Chicago, IL: University of Chicago, Department of Education, 1988), and A. S. Bryk, S. W. Raudenbush, M. Seltzer, and R. Congdon, *HLM/2L* (Chicago, IL: Statistical Software Incorporated, 1992). A preliminary version of the "C" version of HLM/2L was used for this analysis.

spreadsheet results were used to validate its accuracy. HLM2PV is available as part of a special version of HLM/2L.[55]

For each HLM model, HLM2PV runs the HLM/2L program on each of the five (or the number specified) plausible values internally, and produces the average value of each parameter and the correct standard errors. Although the user seems to be producing one estimate, actually the five HLM estimates from the five plausible values are produced, and their average and measurement error are calculated correctly. Thus, using this procedure ensures an accurate treatment of plausible value data and also saves a tremendous amount of calculation time.

The output of HLM2PV is similar to the HLM/2L program output, except that all of the estimates are averaged over estimates derived from each of the five plausible values. The following HLM parameter estimates are averaged by HLM2PV: the Gammas, the parameter variances (Tau), the Chi-square tests of Tau, the Chi-square values for the test of the homogeneity of the level-1 variance, the Chi-square values for the comparison of the deviance statistics, and the reliabilities. The standard errors of the Gamma are calculated based on the formula above, and the probabilities of any Chi-square values are calculated with the equation used in HLM/2L.

D. Interpreting HLM results: HLM statistics

Overview

This section provides an introduction to the statistics produced by an HLM analysis that are used in this report. Besides the Betas and Gammas, other statistics are helpful in interpreting the within-school parameters and the between-school models. For each of the random within-school parameters in each model (in this study, the intercept, gender, and race-ethnicity), HLM/2L provides the parameter variance, called *Tau*, a test of whether Tau is greater than 0, and the reliability, the percentage of the total variance around each parameter that is represented by parameter variance. In addition, R^{2*}, a measure of how well each model explains the parameter variance, can be calculated. R^{2*} is similar to a linear regression R^2, in that it represents the proportion of the original parameter variance that was explained by a particular between-school model. While the significance of the parameter variance is usually tested with a Chi-square test, the deviance statistic provides a more comprehensive test, and there are certain circumstances when a deviance statistic should be used instead of the Chi-square test. This is discussed in chapter six.

In this report, the Gammas, the reliabilities, the parameter variance (Tau), the test for whether Tau is greater than 0, and the R^{2*} are presented and discussed in the results chapters. Tables A1–A36 in Appendix A are the supporting tables for the HLM results presented in chapters three through five. These tables include the Gammas, the standard errors of the Gammas, the *t* value and significance of the Gammas, the reliability of the parameters, the actual parameter variance, or Tau, still present after each model has been run, the degrees of freedom at the school level for each between-school model, and an estimate of the probability that Tau is greater than 0 given those degrees of freedom. This section explains these statistics in greater detail.

---

[55] A. S. Bryk, S. W. Raudenbush, M. Seltzer, and R. Congdon, *Hierarchical Linear Modeling with the HLM/2L and HLM/3L Programs* (Chicago, IL: Statistical Software Incorporated, 1994)

Gammas and standard errors

The Gammas and their standard errors were calculated using the procedure discussed in the previous "Estimation theory and practice" section. Each Gamma is the average of the five Gammas from five separate HLM analyses, using the five plausible values of achievement. Each standard error is the average of the five standard errors from the five Gammas, plus the standard error between the five Gammas. This allowance for measurement error thus increased the standard errors over those obtained for just one plausible value and made it harder for the school effects to be significant. While this limited the number of significant school effects, it lent greater confidence to the results that were significant.

Significance tests on Gammas

Significance was calculated for each Gamma with a *t* value, which was the value of the Gamma divided by its standard error. The probability of this *t* value being larger than 0 was determined with a two-tailed test of significance, using the alpha levels of .05 and .01 for each Gamma. It is possible that since so many parameter estimates were made in each analysis, lower alpha levels could be used to prevent the buildup of Type I error. This procedure was not followed because other HLM studies have not done so in the past and because this was an exploratory study. However, the issue of appropriate significance tests and the meaning of significant Gammas is one that HLM researchers need to discuss. Bryk recommends just using the results as a guideline for further research, rather than for definite answers.[56]

Parameter variance

Parameter variance, or Tau, is an estimate of the actual non-sampling variation between schools around the parameters of the intercept and the gender, race-ethnicity, and SES coefficients in the within-school equations. The parameter variance usually changes between models. It is highest in the unconditional within-school models, where it indicates how much variance there is around each of the four parameters before any between-school variables are taken into account. The purpose of the between-school models is to identify school-level variables that explain, or reduce, this parameter variance, and thus explain school variations in average achievement and differences in achievement by gender and race-ethnicity.

If the parameter variance is 0, as indicated by a Chi-square or deviance test (see chapter six), either in the within-school models or after any between-school models, then there may be no more parameter variance to explain. These tests are commonly used in HLM analysis to decide if more variables need to be added to the model if there is no more variation or if none was present at the beginning, between-school models or more between-school variables are not needed to explain it. However, since this analysis tested variables in separate theoretical groups rather than by hierarchically entering them into one large equation, these tests were not used to determine whether a model was needed or what variables should be added. However, the average of the probabilities of the Chi-square

---

[56]Personal communication, July 6, 1991.

tests is presented so that the reader can interpret the levels of parameter variance before and after the between-school models. In addition, several models were explored using the deviance statistic to see how much of a difference it made in determining whether the parameter variance was significant. See chapter six for further details.

R^{2*}, or proportion of parameter variance explained

If there is still parameter variance to explain, a measure of how well each model explains the parameter variance is the R^{2*}. It is similar to a linear regression R^2 in that it represents the proportion of the original parameter variance that was explained by a particular between-school model. To obtain the R^{2*} for a parameter in a between-school model, the difference between the original parameter variance in the unconditional model and the parameter variance left from each conditional between-school model is divided by the original parameter variance. While knowing the R^{2*} is very important information about the models, the reliability provides additional perspective on the amount of parameter variance there is to explain.

Reliability

In HLM, reliability refers to the percentage of the total variance around each parameter that is parameter variance. The total variance of each parameter consists of both parameter variance and sampling variance. Parameter variance is the actual variation between schools around the parameters of the intercept and the gender and race-ethnicity coefficients in the within-school equations. This variation can be explained by the between-school models. However, there is also sampling variance around these parameters, from sampling error within the schools, and this cannot be explained by the between-school model because it is essentially error. Reliability thus indicates how much of the total variance can be explained by the between-school models.

Special statistical and interpretive procedures for an exploratory study

Due to the exploratory and methodological nature of this study, several types of procedures and interpretations were used that may not have been followed in a theoretical research study. First, although many interesting results were obtained, the emphasis of the discussion was on how those results were produced rather than the meaning of those results. Causal implications especially, common to most school effects studies, were carefully avoided.

Second, a more liberal significance level was used to identify variables that were on the verge of becoming significant in these models. A significance level of .10 identified variables that were "almost significant," and these variables were mentioned in the results chapters. In addition, if these variables were included in several models, they were monitored to see how their significance changed across models. The purpose of this procedure was to illustrate the sensitivity of the models to variable specification. However, in the summary of results, only variables significant at the .05 level or higher were discussed.

Third, some HLM models were estimated on the gender and race-ethnicity coefficients as dependent variables when they had no parameter variance. Since a lack of parameter variance indicates a lack of variance to model, usually these variables are fixed, or at least not modeled. However, the strength of this study—the large number of models, grades, and subjects, which were each run five times to account for the NAEP plausible values—made it impractical to modify each model according to the best specifications. Chapter six, the exploratory chapter, seeks to examine and modify several models more closely in that way.

In the main part of the study, models were developed theoretically, tested with many preliminary runs on one plausible value to find the best variables for most grades, and estimated in final runs with five plausible values, using the same variables regardless of variance. A benefit to this uniform approach to model specification was to illustrate how the HLM statistics behaved with a variety of types of variables and variability.

Chapter III

School Correlates of Average Mathematics Achievement

Overview

The next three chapters present the school-level results of the HLM models tested in this study. Although each HLM model includes the three school-level equations that predict each of the randomly varying student-level parameters, these equations are presented separately in these chapters for purposes of clarity. The results of all three chapters are discussed in chapter seven.

This chapter presents the school-level equations that predict the intercept, or average achievement in each school. Chapter four presents the school-level equations that predict the gender gap, while chapter five presents the school-level equations that predict the race-ethnicity gap in each school. In addition, chapter six presents and discusses the results of an exploratory analysis that examined more closely the parameter variance of the gender gap and race-ethnicity gap equations.

The average achievement within schools is expressed in the unconditional equation 2.2, and the models estimated to predict average achievement between schools are expressed in the conditional equation 2.6 (table 3.1). The intercept equation in the between-school HLM models (equation 2.6) tests the association of various school characteristics with the average mathematics and geometry achievement within schools (table 3.1). This chapter presents the average mathematics and geometry achievement levels within schools (tables 3.2 and 3.3) and the results of the six models estimated to predict that average achievement in mathematics and geometry (tables 3.4 to 3.16).

Table 3.1.—Equations related to intercept, or average achievement

Within-school student level equation

$$y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + r_{ij} \qquad (2.1)$$

Between-school school-level equations

Unconditional intercept equation:

$$\beta_{0j} = \gamma_{00} + u_{0j} \qquad (2.2)$$

Conditional intercept equation:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} W_{01j} + \gamma_{02} W_{02j} + \ldots + \gamma_{0m} W_{0mj} + u_{0j} \qquad (2.6)$$

NOTE: For explanation of terms in equation, see chapter two.

A. Within-school models

Tables 3.2 and 3.3 show the results of the unconditional models for mathematics and geometry, respectively. These models provide the average within-school parameters. Although this chapter focuses only on the intercept parameter, all the parameters in the models are shown because they function as control variables. The average achievement scores in math are about 211 in grade 4, 260 in grade 8, and 292 in grade 12 (table 3.2). The average achievement scores in geometry are 213 in grade 4, 258 in grade 8, and 292 in grade 12 (table 3.3). Because the predictors have 0 means due to centering, these achievement levels represent average achievement in each school at the "average" gender, race-ethnicity, SES, and course-taking values, before controlling for these variables.

These averages are slightly different than the means in table 2.4 because they are the average of the school intercepts, or averages, whereas the means in table 2.4 are the achievement averages across all students. In addition, the averages in tables 3.2 and 3.3 are the average of five estimates of the school averages. According to table 2.4, the average achievement scores have standard deviations in math and geometry, respectively, across all students of 29 and 32 in grade 4, 33 and 35 in grade 8, and 34 and 41 in grade 12.

The intercept averages are larger in the higher grades because the scales have been constituted as a continuum from lower to higher grades. NAEP has selected scale anchor levels, which describe the types of abilities a student at each level on the scale would have, and these anchor levels can be used to compare the intercept averages. The NAEP anchor levels are:[57]

200: Simple additive reasoning and problem solving with whole numbers;
300: Simple multiplicative reasoning and two-step problem solving;
350: Reasoning and problem solving involving fractions, decimals, percents, elementary geometric properties, and simple algebraic manipulations; and
400: Reasoning and problem solving involving geometric relationships, algebraic equations, and beginning statistics and probability.

Tables 3.2 and 3.3 also include the reliability and the original parameter variance around the intercept, or average achievement level, in this unconditional model. The reliability is the proportion of the total variance that the parameter variance represents. It is this parameter variance in the unconditional model that the subsequent HLM models will seek to explain. The rest of the total variance is sampling error, and cannot be explained. The reliability changes little between models, as can be seen in the later tables. However, the parameter variance is usually lower in subsequent conditional models, as variables in these models explain at least some of this variance.

In all three grades and in both math and geometry, the reliability on average achievement is around 90 percent, so most of the variance around average achievement is parameter variance and has the potential to be explained. In addition, in each grade and subject, the parameter variance is high and very significant. Thus, the average achievement varies quite a bit between schools, and this variance has the potential to be explained by a correctly-specified conditional model.

---

[57]A. Rogers et al., *National Assessment of Educational Progress: 1990 Secondary-use Data Files User Guide* (Princeton, NJ: Educational Testing Service, Revised-June 1992): 140.

Table 3.2.—Average within-school predictors of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS	(Unit of parameter)	Grade 4	Grade 8	Grade 12
INTERCEPT	(Average Achievement)	211.35**	260.28**	291.98**
GENDER	(1=Female)	-1.46†	-0.63	-2.53**
RACE-ETHNICITY	(1=Afr. Am/Hisp./Nat. Am)	-14.74**	-15.48**	-14.13**
SES	(0=Mean)	6.20**	10.65**	12.36**
TAKING ALGEBRA	(1=If currently taking algebra)		31.66**	
YRS OF CALCULUS	(1=one year)			17.58**
OTHER HLM STATISTICS FOR INTERCEPT PARAMETER ONLY				
Reliability		.90	.94	.93
Parameter variance		235.98**	281.45**	299.05**

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 3.3.—Average within-school predictors of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS	(Unit of parameter)	Grade 4	Grade 8	Grade 12
INTERCEPT	(Average Achievement)	213.13**	258.04**	292.08**
GENDER	(1=Female)	-0.14	-1.23	-6.47**
RACE-ETHNICITY	(1=Afr. Am/Hisp./Nat Am.)	-12.43**	-14.00**	-14.30**
SES	(0=Mean)	4.53**	8.93**	7.49**
TAKING ALGEBRA	(1=If currently taking algebra)		27.84**	
YRS OF GEOMETRY	(1=One year)			23.58**
OTHER HLM STATISTICS FOR INTERCEPT PARAMETER ONLY				
Reliability		.88	.90	.93
Parameter variance		252.61**	260.16**	392.68**

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

B. Between-school models

Student body characteristics

Tables 3.4 and 3.5 show the HLM results for the student body characteristics models predicting average achievement.[58] The variables in this model also function as control variables for all later models, so their performance in later models is also described.

In all three grades, the percentage of African-American students in the school and the SES level of students in that grade in the school were significantly associated with average achievement levels between schools in both math and geometry. However, the percentage of Hispanic students in a school was not significantly associated with average math achievement in any grades, although in grade 8 math, it was close to being significant. This pattern remained the same in all grades when other variables were added to this model.

The higher the percentage of African-American students in the school, the lower the average achievement by a similar amount in each grade. For every standard deviation above the average percentage of African-American students in a school, the average achievement was 5 to 7 points lower in math and geometry in that school (tables 3.4 and 3.5). For instance in grade 4, the average percentage of African-American students was 13 percent and the standard deviation was 20 percent, so for every 20 percentage points of African-American students above 13 percent, the average math achievement in a school was 5 points lower than in schools with 13 percent African-American students (table 2.4).[59]

SES levels were also associated with average achievement. The higher the average SES level of the students in each grade and school, the higher the math and geometry achievement level in that grade and school. These effects were more pronounced in the higher grades, and ranged from a low of about 5 achievement points in grade 4 geometry to a high of about 13 achievement points in grade 12 geometry for every standard deviation above the average SES (tables 3.4 and 3.5). The mean of the between-school SES levels was 0, and the standard deviation was about .25 in each grade.(table 2.4) Therefore, for every .25 above 0 in the average SES level of a school, the average achievement in that school was about 5 points higher in grade 4, about 9 points higher in grade 8, and about 13 points higher in grade 12.

As in the unconditional model, the reliability for each grade and subject indicates that most of the variance around the intercept is parameter variance. The amount of parameter variance is less than in the unconditional model, but it is still high and significant for each grade and subject, so there is more variance that can be explained with a fuller model. However, this model has already explained from 27 to 29 percent of the grade four variance, and from 70 to 74 percent of the grade eight and twelve variance. Therefore, different variables need to be found to explain more of the grade four variance in average achievement, while this model already explains most of this variance in grades eight and twelve. However, other variables could still add explanatory power to these grades.

---

[58]The average achievement, gender, and race-ethnicity parameters were random and modeled, while the SES and course-taking parameters were fixed and not modeled. Although all of these parameters were included in the full model, only the equation 2.6, which modeled the intercept parameter, is shown in tables 3.4 to 3.16. The results of equations 2.7 and 2.8, which model the gender and race-ethnicity parameters, are presented and discussed in the next two chapters. For the full model, see the corresponding HLM tables in Appendix A.

Table 3.4.—Student body characteristics predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
INTERCEPT (AVERAGE ACHIEVEMENT)			
Intercept	212.87**	262.96**	292.80**
Percent African-American	-5.00**	-7.11**	-6.55**
Percent Hispanic	-0.97	-1.19†	0.91
Average SES	5.01**	9.31**	11.64**
OTHER HLM STATISTICS			
Reliability	.90	.80	.77
Parameter variance	166.91**	73.84**	76.66**
Proportion of parameter variance explained	.29	.74	.74

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the intercept parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education. National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 3.5.—Student body characteristics predictors of student-level parameters of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
INTERCEPT (AVERAGE ACHIEVEMENT)			
Intercept	214.64**	260.42**	292.97**
Percent African-American	-5.00**	-7.56**	-7.18**
Percent Hispanic	-1.25	-0.82	1.11
Average SES	4.81**	8.25**	13.18**
OTHER HLM STATISTICS			
Reliability	.84	.74	.79
Parameter variance	185.33**	77.77**	112.78**
Proportion of parameter variance explained	.27	.70	.71

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the intercept parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

[59]Table 2.4 lists the unstandardized means and standard deviations of all the between-school variables.

School resources (student, staff, fiscal, and physical)

Tables 3.6 and 3.7 show the HLM results for the school resources models for all three grades, and table 3.8 shows the HLM results for a detailed school resources model for grade 4. For the most part, differences in school resources were not associated with differences in average math or geometry achievement levels between schools. The amount of district instructional funds spent per student, the student/teacher ratio, and the number of students in the school were not related to student achievement in any grade. Reflecting this lack of association of school resource variables with average achievement in schools, the parameter variance and, thus, the proportion of variance explained, was not reduced any further by this model than by the previous school characteristics model.

In the grade 4 detailed model only, the higher the percentage of students in the school using computers in instruction, the higher the average achievement in math and geometry (table 3.8) by about 2 to 3 achievement points for every standard deviation above the average percentage. However, the average number of computers per student was not related to achievement in grade 4.[60] The proportion of parameter variance explained increased only slightly for grade four with this model, from .29 to .31 in mathematics and from .27 to .30 in geometry, indicating that this one significant school resource variable did not add much explanatory power to the school characteristics model.

Table 3.6.—School resources predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
INTERCEPT (AVERAGE ACHIEVEMENT)			
Intercept	213.12**	262.89**	292.55**
Percent African-American	-4.69**	-7.18**	-6.60**
Percent Hispanic	-0.18	-1.38†	0.48
Average SES	5.07**	9.22**	10.98**
School size (number of students)	-1.34	0.04	1.10
Student/teacher ratio	-1.36	0.59	-0.89
District instructional funds/student	-0.68	0.39	0.76
OTHER HLM STATISTICS			
Reliability	.86	.80	.77
Parameter variance	166.48**	76.17**	77.21**
Proportion of parameter variance explained	.29	.73	.74

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the intercept parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

[60]These computer variables could not be tested in grades 8 and 12 due to the large number of missing values. See chapter two for a discussion of missing values.

Table 3.7.—School resources predictors of student-level parameters of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
INTERCEPT (AVERAGE ACHIEVEMENT)			
Intercept	214.96**	260.20**	292.62**
Percent African-American	-4.54**	-7.73**	-7.25**
Percent Hispanic	-0.35	-1.21	0.56
Average SES	4.92**	7.88**	12.12**
School size (number of students)	-1.56	0.31	1.46
Student/teacher ratio	-1.29	0.51	-1.64
District instructional funds/student	-1.09	0.98	0.24
OTHER HLM STATISTICS			
Reliability	.84	.74	.79
Parameter variance	183.43**	79.09**	111.08**
Proportion of parameter variance explained	.27	.70	.72

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the intercept parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 3.8.—Detailed school resources predictors of student-level parameters of grade 4 math and geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Math Grade 4	Geometry Grade 4
INTERCEPT (AVERAGE ACHIEVEMENT)		
Intercept	213.30**	215.23**
Percent African-American	-4.09**	-3.86**
Percent Hispanic	-0.35	-0.60
Average SES	5.07**	4.93**
School size (number of students)	-1.33	-1.40
Student/teacher ratio	-1.58	-1.45
District instructional funds/student	-0.71	-1.10
Computers per student	-0.83	-0.51
Percent use computers as part of math instruction	2.36*	2.52*
OTHER HLM STATISTICS		
Reliability	.86	.83
Parameter variance	162.09**	177.98**
Proportion of parameter variance explained	.31	.30

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the intercept parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Classroom instructional methods

Tables 3.9 and 3.10 show the HLM results for the classroom instructional methods models for all three grades. The amount of time spent in the seven types of instructional methods (nine types in grade 12) was based on student perceptions and averaged across the grade within schools. Several of these methods were associated with different levels of average achievement by school, controlling for percent African-American, percent Hispanic, and average SES levels.

For overall math achievement in each grade, the more time spent doing problems from textbooks in math class, the higher was the average math and geometry achievement. The average school achievement score was 2–3 points higher for every standard deviation above the average amount of time spent working from textbooks (tables 3.9 and 3.10). In addition, in grade 4, the more work in math class that was done with objects (blocks, rulers, and shapes), the higher was the average achievement in math by about 2 points for every standard deviation above average work with objects. However, also in grade 4, the more that calculators were used in math class, the worse students did in math, also by about 2 points for every standard deviation above average work with calculators. Similarly, the more that computers were used in 8th grade math class, the lower the scores of 8th grade students in math achievement by about 2 points for every standard deviation above average use of computers. In grade 12, however, using calculators was associated with higher achievement in math by 3.5 points for every standard deviation above average use of calculators. In addition, in schools where 12th graders more than average time writing proofs in math class, math achievement was higher by about 2 points for every standard deviation above the average time spent writing math proofs.

For the geometry subscale, some similar instructional methods were associated with achievement. Working on problems from textbooks was positively associated with achievement only in grades 4 and 8. As was true for math achievement, computers in grade 8 were associated with lower geometry achievement, while using calculators in grade 12 was associated with higher geometry achievement. These results will be discussed in the discussion chapter.

These classroom instructional methods models, with several significant variables, improved the proportion of variance explained somewhat by several percentage points above the student characteristics models. Still, a significant amount of parameter variance was left to explain.

In addition to the significant variables in these models, there were several methods that were almost significant, with a significance level of between .055 and .10. At this marginal level of significance, working with objects was positively associated with geometry achievement in grades 4 and 8, but negatively associated in grade 12. In addition, in grade 8, taking math tests was negatively associated with achievement. In grade 12, schools in which 12th graders spent more time writing math proofs averaged higher geometry achievement, while those that spent more time formulating their own math problems averaged lower achievement. If other methods or other variables were added to this model, these methods could be examined to see if any become significant at a level of less than .05 or if they become less significant. However, without a comparison model, there is no way to tell whether these results are true relationships or random non-significant associations.

Table 3.9.—Classroom instructional methods predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
INTERCEPT (AVERAGE ACHIEVEMENT)			
Intercept	213.04**	262.78**	292.55**
Percent African-American	-4.70**	-6.83**	-6.31**
Percent Hispanic	-0.48	-0.78	0.22
Average SES	4.83**	9.21**	8.60**
Average time spent on:			
Working in small groups	-1.31	-1.02	0.91
Working with objects	1.96*	0.64	-1.47
Doing problems on worksheets	0.05	1.70	-0.75
Doing problems from textbook	2.74**	2.48**	3.03*
Taking math tests	-1.30	-1.35	-0.63
Using computer	0.96	-2.16*	-0.45
Using calculator	-2.12*	1.26	3.53**
Writing math proofs			1.99*
Formulating own math problems			-1.52
OTHER HLM STATISTICS			
Reliability	.85	.78	.72
Parameter variance	152.19**	65.99**	58.93**
Proportion of parameter variance explained	.36	.76	.80

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the intercept parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 3.10.—Classroom instructional methods predictors of student-level parameters of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
INTERCEPT (AVERAGE ACHIEVEMENT)			
Intercept	214.81**	260.24**	292.71**
Percent African-American	-4.68**	-7.09**	-6.79**
Percent Hispanic	-0.75	-0.45	0.31
Average SES	4.60**	8.11**	9.69**
Average time spent on:			
Working in small groups	-1.24	-1.00	-0.53
Working with objects	2.02†	1.73†	-2.00†
Doing problems on worksheets	0.22	2.12	-0.93
Doing problems from textbook	2.96**	2.07*	2.52
Taking math tests	-1.33	-2.01†	0.15
Using computer	0.96	-2.03*	-0.32
Using calculator	-1.80†	1.47	4.27**
Writing math proofs			2.32†
Formulating own math problems			-2.48†
OTHER HLM STATISTICS			
Reliability	.83	.71	.76
Parameter variance	171.23**	66.79**	90.97**
Proportion of parameter variance explained	.32	.74	.77

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the intercept parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade ' 12th Grade Students, Restricted-Use Data Base.

School climate models

In these school climate models, three types of school climate equations were estimated. The first model focused on attitudes held by the students about math, about themselves and math, and about females and math averaged across the grade within each school. Using these attitudes as control variables, the second model added student behavior and physical safety variables to the equation. The third model also used the attitude variables as controls, and tested the association of academic expectations with achievement. These school climate models will be used to illustrate the sensitivity of HLM equations to variable specification in the models.

Mathematics Attitudes

Tables 3.11 and 3.12 show the HLM results for the math attitudes models for all three grades. Student attitudes alone, controlling for percent African-American, percent Hispanic, and average SES levels, were associated with achievement in grades 8 and 12. Schools in which more 8th grade students liked math and felt they were good at it had higher achievement in 8th grade math overall by about 2.5 points (table 3.11). In grade 12, feeling positive about math was strongly associated with higher achievement in math and geometry (table 3.12). Average achievement in these schools was 7 to 8 points higher for every standard deviation above the average attitude about math. However, the more 12th grade students who believed that math was useful, the lower the average math and geometry achievement in the school by about 3 points in math and 5 points in geometry (tables 3.11 and 3.12). These results will be discussed in the discussion chapter.

In terms of proportion of parameter variance explained, this basic student climate/math attitudes model did not do much better than any of the previous models in grades 4 and 8. For these grades, the explanatory power of the school characteristics model did not improve by more than two percentage points for grades 4 and 8 in overall math and geometry achievement. However, for grade 12, the proportion of variance explained increased by 8 percentage points for both overall math and geometry. Therefore, this model had the most salience for grade 12 math and geometry achievement above the student body characteristics.

For all grades, there was still significant amount of parameter variance left to explain, so the subsequent school climate models, which were built upon this model, were tested to see if they could improve the explanatory power for any of the grades.

In addition, in this first school climate model, there were several variables that were almost significant and that can be examined further in the subsequent school climate models. In grade 4, schools in which more students held female-positive attitudes about math had higher achievement in both math and geometry (tables 3.11 and 3.12). In the grade 8 overall math equation, schools in which more students held female-positive attitudes about math were more likely to have higher average 8th grade math achievement (table 3.11). In addition, schools with higher percentages of Hispanic students had slightly lower 8th grade math achievement (table 3.11). And in grade 12, schools with higher percentages of Hispanic students had slightly higher 12th grade math achievement (table 3.11). Because the subsequent school climate models build upon this model, these variables can be examined in these models to see if their significance increases or decreases. We then might begin to determine whether these results are true relationships or random non-significant associations. The performance of these and other variables will demonstrate the sensitivity of these models to variable specification.

Table 3.11.—Student climate (math attitudes) predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
INTERCEPT (AVERAGE ACHIEVEMENT)			
Intercept	213.02**	263.09**	292.59**
Percent African-American	-4.59**	-7.43**	-7.17**
Percent Hispanic	-0.90	-1.13†	1.03
Average SES	3.71**	9.14**	10.79**
Students feel math is useful	1.44	-1.25	-3.33**
Students enjoy and feel competent in math	1.26	2.52*	6.66**
Students disagree that math is more for boys	1.92†	1.49†	-0.06
OTHER HLM STATISTICS			
Reliability	.86	.79	.71
Parameter variance	163.15**	69.11**	54.69**
Proportion of parameter variance explained	.31	.75	.82

[1] All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the intercept parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 3.12.—Student climate (math attitudes) predictors of student-level parameters of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
INTERCEPT (AVERAGE ACHIEVEMENT)			
Intercept	214.81**	260.55**	292.70**
Percent African-American	-4.50**	-7.76**	-7.78**
Percent Hispanic	-1.20	-0.79	1.31†
Average SES	3.54**	8.11**	12.09**
Students feel math is useful	1.73	-1.51	-4.49**
Students enjoy and feel competent in math	0.60	1.50	7.58**
Students disagree that math is more for boys	1.95†	1.28	-0.28
OTHER HLM STATISTICS			
Reliability	.84	.73	.74
Parameter variance	181.91**	76.14	82.99*
Proportion of parameter variance explained	.28	.71	.79

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the intercept parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Student Behavior and Safety

Tables 3.13 and 3.14 show the HLM results for the student behavior and safety models for all three grades. None of the student behavior or physical safety problems were associated with achievement differences in grade 4 or 8 (tables 3.13 and 3.14). However, by controlling for problems in the school and the percent of students enrolled all year, the female-positive attitudes that had been almost significant and positively associated with math achievement in grade 4 became significant (table 3.13). The other variables that were almost significant in grade 4 and 8—female positive attitudes in grade 4 geometry and grade 8 math and percent Hispanic in grade 8 math and geometry—remained at the same marginal level of significance with similar gamma values and directions (tables 3.13 and 3.14).

In grade 12, one of the student behavior and physical safety problems was associated with math and geometry achievement. Schools where more students reported that other students disrupted class, had a lower average achievement in math and geometry by about 3 points for every standard deviation above the average of class disruptions (tables 3.13 and 3.14). By controlling for these problems, the previous association of percent Hispanic with geometry achievement dropped from almost significant to not significant at all.

Not surprisingly, this model did not improve the proportion of parameter variance explained above that of the previous school climate model for grades 4 or 8 at all, and only increased the explanatory power for grade 12 by 3 percentage points. All grades still had significant amounts of parameter variance left to explain.

The next school climate model did not build upon this model, but replaced the student behavior and safety variables with predictors about academic expectations, leaving the basic math attitudes in the model.

Table 3.13.—Student climate (math attitudes and student behavior and safety) predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
INTERCEPT (AVERAGE ACHIEVEMENT)			
Intercept	213.61**	263.06**	293.06**
Percent African-American	-4.27**	-7.75**	-6.25**
Percent Hispanic	-0.79	-1.24†	1.03
Average SES	3.59**	9.26**	9.98**
Students feel math is useful	1.47	-1.25	-3.22**
Students enjoy and feel competent in math	1.47	2.59*	6.42**
Students disagree that math is more for boys	2.04*	1.68†	-0.30
Index of problems in the school	-0.59		
Percent students enrolled all year	0.64		
Absenteeism in grade		-0.58	0.84
Students feel classes often disrupted		-0.58	-2.84**
Students feel unsafe at school		1.35	-1.36
OTHER HLM STATISTICS			
Reliability	.86	.79	.67
Parameter variance	161.98**	69.32**	45.73**
Proportion of parameter variance explained	.31	.75	.85

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the intercept parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 3.14.—Student climate (math attitudes and student behavior and safety) predictors of student-level parameters of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
INTERCEPT (AVERAGE ACHIEVEMENT)			
Intercept	215.58**	260.50**	293.21**
Percent African-American	-4.06**	8.06**	-6.78**
Percent Hispanic	-1.05	-0.89	1.26
Average SES	3.38**	8.22**	11.28**
Students feel math is useful	1.78	-1.35	-4.38**
Students enjoy and feel competent in math	0.89	1.61	7.32**
Students disagree that math is more for boys	2.11†	1.45	-0.55
Index of problems in the school	-0.78		
Percent students enrolled all year	0.83		
Absenteeism in grade		-0.45	1.19
Students feel classes often disrupted		-0.37	-3.22**
Students feel unsafe at school		1.16	-1.33
OTHER HLM STATISTICS			
Reliability	.83	.74	.71
Parameter variance	178.10**	76.91**	71.14**
Proportion of parameter variance explained	.29	.70	.82

[1] All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the intercept parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Academic Expectations

Tables 3.15 and 3.16 show the HLM results for the academic expectations models for all three grades. Each grade had different variables measuring academic expectations, but academic expectation variables in each grade were associated with achievement only in grades 8 and 12.

In grade 4, neither the amount of math instruction per week in the school nor whether math was a priority in the school were related to average math or geometry achievement (tables 3.15 and 3.16). However, controlling for these variables made female-positive attitudes significantly associated with both math and geometry achievement in grade 4.

The female-positive attitudes variable wavered between almost significant and significant for math and geometry in these three school climate models (tables 3.11-3.16). The fact that three variables, two related to math emphasis in the school, were needed to push this variable to significance showed the possible weakness of this association. This variable would need to be tested in models with a variety of other variables in order to determine how robust this relationship was.

Similarly in grade 4, in the last school climate model, the math attitudes variable of students feeling math is useful went from being not at all significant to being almost significant and positively associated with geometry achievement. Given that this was not at all significantly related to geometry achievement in the earlier models, the robustness of this association is also suspect.

In grade 8, the percentage of 8th graders taking algebra was the only academic expectation variable. Schools with higher percentages of 8th graders taking algebra had higher average achievement in math and geometry by about 3 points for every standard deviation above the average of percentages taking algebra (tables 3.15 and 3.16).

By controlling for the percentages of 8th graders taking algebra, two variables became significant in grade 8 for the first time. The association of female-positive attitudes with higher average math achievement went from almost significant to significant for math and, for the first time, became almost significant for geometry (tables 3.11 to 3.16). In addition, the almost significant negative association of percentage of Hispanic students with math achievement became significant, while the non-association of this variable in geometry became negatively and almost significant (tables 3.11 to 3.16). However, as in grade 4, these variables would need to be tested in models with a variety of other variables in order to determine how robust these relationships were.

In grade 12, both academic expectation variables—the percentage of 12th graders on the academic/college prep track and the average number of years of calculus taken by 12th grade—were positively associated with achievement. The more 12th graders on the academic/college prep track, and the higher amount of calculus taken, the higher the school's average grade 12 achievement in both math and geometry by about 4 to 5 points and 2 to 3 points, respectively, for every standard deviation above the average of these academic expectations variables (tables 3.15 and 3.16). As in the previous school climate model, by controlling for these academic expectations, the association of percent Hispanic with geometry achievement in the first school climate model dropped from almost significant to not significant at all.

After this last model, the parameter variance in all grades and in both math and geometry remained high and significantly different from 0, indicating that more models or variables were needed to explain this variance. The proportion of variance explained rose only slightly for all three grades in math and geometry, although for grades 8 and 12 this model had the highest explanatory power of the six models tested. This model explained 78 and 72 percent of the variance in grade 8 math and geometry, respectively, and it explained 86 and 83 percent of the variance in grade 12 math and geometry, respectively. However, it explained only 32 and 30 percent of the variance in grade 4 math and geometry, respectively. Therefore, while a large portion of the grade 8 and 12 variance in average achievement had been explained, the models were for the most part incorrectly specified to explain the grade 4 variance in average achievement.

Table 3.15.—School climate (math attitudes and academic expectations) predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
INTERCEPT (AVERAGE ACHIEVEMENT)			
Intercept	213.75**	262.82**	292.38**
Percent African-American	-4.15**	-7.49**	-6.90**
Percent Hispanic	-0.83	-1.75*	0.43
Average SES	3.75**	7.86**	8.22**
Students feel math is useful	1.55	-0.69	-1.90*
Students enjoy and feel competent in math	1.43	2.34*	4.92**
Students disagree that math is more for boys	2.07*	1.95*	-0.21
Amount of instruction in math	-0.24		
Math identified as a special priority	-0.53		
Percent of 8th grade students taking algebra		3.07**	
Percent of students on academic/college prep			3.84**
Mean years 12th graders have taken calculus			2.12**
OTHER HLM STATISTICS			
Reliability	.86	.77	.65
Parameter variance	161.12**	62.19**	42.97**
Proportion of parameter variance explained	.32	.78	.86

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the intercept parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 3.16.—School climate (math attitudes and academic expectations) predictors of student-level parameters of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
INTERCEPT (AVERAGE ACHIEVEMENT)			
Intercept	215.72**	260.33**	292.42**
Percent African-American	-3.99**	-7.81**	-7.45**
Percent Hispanic	-1.12	-1.31†	0.60
Average SES	3.64**	7.03**	9.01**
Students feel math is useful	1.90†	-0.84	-2.73*
Students enjoy and feel competent in math	0.81	1.36	5.47**
Students disagree that math is more for boys	2.17*	1.65†	-0.46
Amount of instruction in math	-0.54		
Math identified as a special priority	-0.03		
Percent of 8th grade students taking algebra		2.54**	
Percent of students on academic/college prep			4.71**
Mean years 12th graders have taken calculus			2.51**
OTHER HLM STATISTICS			
Reliability	.83	.73	.69
Parameter variance	176.76**	72.78**	65.98**
Proportion of parameter variance explained	.30	.72	.83

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the intercept parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

C. Summary

Significant predictors

Grade 4

The following school characteristics were identified as correlates of average math and geometry achievement in grade 4. The percentage of African-American students in a school was associated with lower achievement while the SES level of a school was associated with higher achievement. Controlling for these student body characteristics, schools in which higher percentages of students were using computers had higher math and geometry achievement. In the classroom, the more time spent doing problems from textbooks the higher the average math and geometry achievement, while the more time spent working with objects, the higher the average geometry achievement in grade 4. However, the use of calculators in the 4th grade classrooms was associated with lower average math achievement.

School climate also made a difference for 4th grade achievement. When controlling for math academic expectations, schools in which more 4th grade students felt positively about girls and math had higher 4th grade math and geometry achievement, and when controlling for problems in the school and the percentage of full-year students, schools in which more 4th grade students felt positively about girls and math had higher 4th grade math achievement.

Grade 8

The following school characteristics were identified as correlates of average math and geometry achievement in grade 8. As in grade 4, the percentage of African-American students in a school was associated with lower math and geometry achievement while the SES level of a school was associated with higher achievement. Controlling for these student body characteristics, schools in which students in 8th grade classrooms were spending more than average amounts of time doing problems from textbooks, the higher the average math and geometry achievement in grade 8. However, the more time spent working with computers in 8th grade classes, the lower the average math and geometry achievement.

Schools in which more 8th grade students liked math and felt they were good at it had higher average achievement in 8th grade math. Finally, schools with higher percentages of 8th graders taking algebra had higher average achievement in math and geometry. And, when the percentage taking algebra was controlled for, schools in which more 8th grade students held female-positive attitudes about math also had higher average math achievement.

Grade 12

The following school characteristics were identified as correlates of average math and geometry achievement in grade 12. As in grades 4 and 8, the percentage of African-American students in a school was associated with lower achievement while the SES level of a school was associated with higher achievement.

Controlling for these student body characteristics, the following additional results were found. As in grades 4 and 8, schools in which students in 12th grade classrooms were spending more than average amounts of time doing problems from textbooks averaged higher math achievement in grade 12. In addition, 12th graders also averaged higher math and geometry achievement in schools where they were more likely to use calculators and average math achievement in schools where they were more likely to write math proofs in class.

Several school climate measures were associated with 12th grade achievement. Feeling positive about math was strongly associated with higher achievement in math and geometry. However, the more 12th grade students who believed that math was useful, the lower the average math and geometry achievement. Grade 12 registered the only student behavior and safety issue. In schools where 12th graders felt that their classes were disrupted by other students, math and geometry achievement was lower. Finally, the more 12th graders on the academic/college prep track, and the higher average amount of calculus taken by the 12th grade, the higher the school's average grade 12 achievement in both math and geometry.

Proportion of parameter variance explained and reliability

The six estimated HLM models identified some school characteristics that were significantly associated with the average math and geometry achievement in each grade. In addition, the HLM program also provided estimates of the amount of parameter variance around average achievement that was left to explain after each model had been estimated. This variance was compared to the amount of parameter variance that existed in the unconditional model before any models were estimated and presented as a proportion of parameter variance that was explained by each model.

Tables 3.17 and 3.18 present the proportion of parameter variance that was explained by each model within each grade. Each model has the potential to explain 1.00, or 100 percent of the parameter variance. Since these six models were tested separately, with only a few common variables, for the most part these proportions should be seen as a comparison of the explanatory power of the different models, rather than as a cumulative or additive measure across the rows of the table. However, since all models did contain the student body variables, all models can be compared to the first model to see if other variables added any explanatory power. In addition, since all three school climate models contained the math attitudes variables, the last two models can be compared to each other to see if either add any explanatory power to the math attitudes model.

For the intercept, or average achievement parameter in grades 8 and 12, the school climate models, especially the academic expectations model, worked best, explaining between 70 and 78 percent of this variance in grade 8 and between 82 to 86 percent of this variance in grade 12. For grade 4, the classroom instructional methods model explained the most of the parameter variance, but even in this model it only explained between 32 and 36 percent. Thus, while the classroom instructional methods variables were significantly associated with average achievement in grade 4, other variables need to be tested to explain more of the variance. However, for grades 8 and 12, these models explained a good majority of the variance.

In the intercept model, reliability was about 90 percent for each grade. While the proportion of parameter variance explained rose in grades 8 and 12, the reliability fell from around 90 percent to around 65 to 75 percent, indicating that as more variables were added to the equation, more sampling error was brought in. However, in the grade 4 models, reliability remained above 80 percent, despite the additional variables.

Table 3.17.—Proportion of parameter variance explained by each HLM model for the intercept parameter of math achievement, grades 4, 8, and 12

Parameter	HLM Models			School Climate Models		
	Student Body Characteristics	School Resources	Classroom Instructional Methods	Math Attitudes	Behavior & Safety & Attitudes	Academic Expectations & Attitudes
Grade 4 Mathematics INTERCEPT	0.29	0.29	0.36	0.31	0.31	0.32
Grade 8 Mathematics INTERCEPT	0.74	0.73	0.76	0.75	0.75	0.78
Grade 12 Mathematics INTERCEPT	0.74	0.74	0.80	0.82	0.85	0.86

NOTE: These proportions are calculated from *average* Tau values (averaged across the five plausible values). Negative proportions due to sampling variation have been set to 0.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 3.18.—Proportion of parameter variance explained by each HLM model for the intercept parameter for geometry achievement, grades 4, 8, and 12

Parameter	HLM Models			School Climate Models		
	Student Body Characteristics	School Resources	Classroom Instructional Methods	Math Attitudes	Behavior & Safety & Attitudes	Academic Expectations & Attitudes
Grade 4 Geometry INTERCEPT	0.27	0.27	0.32	0.28	0.29	0.30
Grade 8 Geometry INTERCEPT	0.70	0.70	0.74	0.71	0.70	0.72
Grade 12 Geometry INTERCEPT	0.71	0.72	0.77	0.79	0.82	0.83

NOTE: These proportions are calculated from *average* Tau values (averaged across the five plausible values). Negative proportions due to sampling variation have been set to 0.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Chapter IV

School Correlates of Gender Differences in Mathematics Achievement

Overview

This chapter presents the school-level equations that predict the "gender gap," or the overall gender differences in achievement within schools. The average gender gap is expressed in the unconditional equation 2.3, and the models estimated to predict the gender gap between schools are expressed in the conditional equation 2.7 (table 4.1). The gender equation in the between-school HLM models (equation 2.7) tests the association of various school characteristics with the mathematics and geometry achievement gaps within schools between female students and male students (table 4.1). This chapter presents the average gender gap in mathematics and geometry within schools (tables 4.2 and 4.3) and the results of the six models estimated to predict the gender gap in mathematics and geometry (tables 4.4 to 4.16).

Table 4.1.—Equations related to gender beta coefficient, or the gender gap

Within-school student level equation

$$y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + r_{ij} \qquad (2.1)$$

Between-school school-level equations

Unconditional gender equation:

$$\beta_{1j} = \gamma_{10} + u_{1j} \qquad (2.3)$$

Conditional gender equation:

$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_{11j} + \gamma_{12}W_{12j} + \dots + \gamma_{1m}W_{1mj} + u_{1j} \qquad (2.7)$$

NOTE: For explanation of terms in equation, see chapter two.

A. Within-school models

Overall gender differences within schools

Tables 4.2 and 4.3 show the results of the unconditional models for the gender coefficient predicting math and geometry achievement, respectively. These models provide the average gender differences within schools (equation 2.3) for math and geometry. Although this chapter focuses only on the gender parameter, all the parameters in the models are shown because they function as control variables.

In grades 4 and 8 math and geometry, there were no average within-school gender differences in achievement in 1990, controlling for race-ethnicity, SES, and taking algebra in grade 8 (tables 4.2 and 4.3). However, in grade 4 math, the gender gap was almost significant (table 4.2). Within schools in grade 12, females averaged 2.5 points lower than males in math, and 6.5 points lower than males in geometry, controlling for race-ethnicity, SES, and for semesters of calculus or geometry, respectively (tables 4.2 and 4.3).[61]

Tables 4.2 and 4.3 also include the reliability and the original parameter variance around the gender gap in this unconditional model. In all three grades and in both math and geometry, the reliability on the gender coefficient varied between 15 and 20 percent, so less than one fifth of the variance around the gender gap is parameter variance and has the potential to be explained.

In grades 4 and 8, there was some variation among schools on gender differences in math achievement as indicated by significant parameter variances (table 4.2). That is, in some schools, females averaged higher achievement scores than males, and in other schools, males averaged higher scores. In addition, the size of the gap between females and males varied between schools. In addition in grade 12, the parameter variance was almost significant, so there was some indication of variation in this gap as well. In all grades, this variation was modeled by school-level equations that predicted which types of schools had larger or smaller gaps between males and females in mathematics achievement.

However, there was little variation among schools on gender differences in geometry achievement. Thus in grades 4 and 8, most schools averaged no gender gap in geometry (table 4.3). In addition, in grade 12, while females achieved significantly lower than males on average, there was still no significant parameter variance between schools in these differences, that is, a similar gap appeared in most schools (table 4.3). Therefore, for geometry, a model predicting which types of schools had a larger or smaller gender gap might not find enough variation in the gap to make those predictions.

Despite this lack of variation in the gender gap in geometry, the decision was made to model these gender coefficients for two reasons. First, since these models were developed theoretically, it was important to be consistent with the gender gap models for math. Second, estimating these models would illustrate how the school-level variables behave when predicting significant and non-significant student-level coefficients with no variance. The results of the school-level models are reported in the next section.

[61]For models without coursetaking in the within-school equations in grades 8 and 12, see chapter six.

Table 4.2.—Average within-school gender coefficients predicting math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS	(Unit of parameter)	Grade 4	Grade 8	Grade 12
INTERCEPT	(Average Achievement)	211.35**	260.28**	291.98**
GENDER	(1=Female)	-1.46†	-0.63	-2.53**
RACE-ETHNICITY	(1=Afr. Am/Hisp./Nat. Am)	-14.74**	-15.48**	-14.13**
SES	(0=Mean)	6.20**	10.65**	12.36**
TAKING ALGEBRA	(1=If currently taking algebra)		31.66**	
YRS OF CALCULUS	(1=one year)			17.58**
OTHER HLM STATISTICS FOR GENDER PARAMETER ONLY				
Reliability		.15	.21	.18
Parameter variance		23.03*	22.33*	24.18†

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 4.3.—Average within-school gender coefficients predicting geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS	(Unit of parameter)	Grade 4	Grade 8	Grade 12
INTERCEPT	(Average Achievement)	213.13**	258.04**	292.08**
GENDER	(1=Female)	-0.14	-1.23	-6.47**
RACE-ETHNICITY	(1=Afr. Am/Hisp./Nat. Am)	-12.48**	-14.00**	-14.30**
SES	(0=Mean)	4.53**	8.93**	7.49**
TAKING ALGEBRA	(1=If currently taking algebra)		27.84**	
YRS OF GEOMETRY	(1=One year)			23.58**
OTHER HLM STATISTICS FOR GENDER PARAMETER ONLY				
Reliability		.18	.17	.20
Parameter variance		40.06	26.51	35.61

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

B. Between-school models

Predictors of gender differences within schools

Overview

Tables 4.4 to 4.16 show the HLM results for the six models (based on equation 2.7) used to predict the gender gap in math and geometry.[62] This section highlights the school-level variables in each model that were found to be significantly related to the gender gap. Table 4.17 summarizes the school-level variables that were significantly associated with the gender gap in math and geometry achievement and shows the number of achievement points in the gap associated with a change in each variable by one standard deviation. Tables 4.18 and 4.19 display the proportion of parameter variance explained for the math and geometry gender gap by each model.

Specific Models

In all but one of the six models, no variables were significantly associated with the gender gap in either mathematics or geometry (tables 4.6 to 4.17). Only in the classroom instructional methods model was one variable significantly related to the gender gap in geometry achievement. However, in the school resources, classroom instructional methods, and school climate models, there were variables that were almost significant (tables 4.6 to 4.17).

In the student body characteristics models, no variables were related to the gender gap in mathematics or geometry, despite the presence of parameter variance in the gender gap in grade 4 and 8 mathematics. Consequently, the proportion of parameter variance explained by this model was close to zero, and there was still parameter variance to explain in the gender gap in mathematics in grades 4 and 8. The parameter variance in the gender gap in grade 12 mathematics remained almost significant. It was not surprising that no variables were significant for geometry since the lack of parameter variance in the gender gap in geometry remained. However, the parameter variance in grade 12 was now almost significant.

Despite their lack of association with the gender gap, the student body characteristics variables were included in all subsequent models as control variables in order to maintain consistent theoretical models and to use the centering method correctly. Since the gender, race-ethnicity, SES, and course-taking variables were centered at the within-school level, each of the school-level models had to include the means of race-ethnicity and SES. The gender mean was not necessary because it was a constant. The course-taking means were only included in the model in which they were of interest

[62]The average achievement, gender, and race-ethnicity parameters were random and modeled, while the SES and course-taking parameters were fixed and not modeled. Although all of these parameters were included in the full HLM model, only equation 2.7, which modeled the gender parameter, is shown in this chapter. For the full model, see the corresponding HLM tables in Appendix A.

Table 4.4.—Student body characteristics predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
GENDER BETA COEFFICIENT			
Intercept	-1.39†	-0.60	-2.66**
Percent African-American	-0.24	-1.40	-0.02
Percent Hispanic	0.14	-1.12	-0.55
Average SES	1.39	-1.51	0.30
OTHER HLM STATISTICS			
Reliability	.15	.20	.19
Parameter variance	22.37*	21.92*	25.83†
Proportion of parameter variance explained	.03	.02	.00

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the gender parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 4.5.—Student body characteristics predictors of student-level parameters of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
GENDER BETA COEFFICIENT			
Intercept	-0.04	-1.18	-6.67**
Percent African-American	-0.25	-1.20	0.79
Percent Hispanic	0.07	-0.62	-0.74
Average SES	1.54	-0.03	1.35
OTHER HLM STATISTICS			
Reliability	.17	.16	.20
Parameter variance	37.28	29.48	38.28†
Proportion of parameter variance explained	.07	.00	.00

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the gender parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

In the school resources model, no variables were significant for the gender gap in either math or geometry, although several were almost significant for math (tables 4.6 to 4.8). In grade 12, larger schools were almost significantly associated with a larger gender gap in math (table 4.6). In grade 4 in the detailed school resources model, larger schools and a larger number of computers per student were almost associated with a smaller gender gap in math achievement (table 4.8).

A non-significant model should have no effect on the parameter variance left to explain, and the parameter variance levels did remain similar to the unconditional model in all grades for the gender gap in both math and geometry. In addition, controlling for school resources, the parameter variance that was almost significant dropped back to non-significance for the gender gap in grade 12 math and geometry.

Table 4.6.—School resources predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
GENDER BETA COEFFICIENT			
Intercept	-1.56†	-0.50	-2.39*
Percent African-American	-0.29	-1.30	0.20
Percent Hispanic	-0.21	-0.82	0.11
Average SES	1.49	-1.23	1.60
School size (number of students)	1.14	-0.46	-1.96†
Student/teacher ratio	0.41	-0.42	1.65
District instructional funds/student	-0.47	-0.56	-1.24
OTHER HLM STATISTICS			
Reliability	.15	.21	.18
Parameter variance	22.24*	22.65*	23.84
Proportion of parameter variance explained	.03	.00	.01

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the gender parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 4.7.—School resources predictors of student-level parameters of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
GENDER BETA COEFFICIENT			
Intercept	-0.10	-1.36	-6.50**
Percent African-American	-0.45	-1.25	0.86
Percent Hispanic	-0.18	-0.80	-0.22
Average SES	1.48	-0.30	2.37
School size (number of students)	0.96	-0.11	-0.72
Student/teacher ratio	-0.04	0.02	0.33
District instructional funds/student	0.33	0.19	-1.94
OTHER HLM STATISTICS			
Reliability	.18	.18	.19
Parameter variance	39.60	28.33	33.42
Proportion of parameter variance explained	.00	.00	.06

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the gender parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 4.8.—Detailed school resources predictors of student-level parameters of grade 4 math and geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Math Grade 4	Geometry Grade 4
GENDER BETA COEFFICIENT		
Intercept	-1.45	-0.18
Percent African-American	-0.62	-0.59
Percent Hispanic	-0.33	-0.26
Average SES	1.65	1.50
School size (number of students)	1.77†	1.30
Student/teacher ratio	0.71	0.13
District instructional funds/student	-0.27	0.38
Computers per student	1.53†	0.98
Percent use computers as part of math instruction	-1.43	-0.64
OTHER HLM STATISTICS		
Reliability	.14	.19
Parameter variance	21.82*	40.68
Proportion of parameter variance explained	.05	.00

[1] All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the gender parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Only one classroom instructional methods variable was associated with the gender gap in either math or geometry. Schools in which computers were used in 12th grade math class for more than average amounts of time had a larger gender gap in geometry than other schools. For every standard deviation above the average amount of computer use in math class, the gender gap in 12th grade was higher by an average of about 3 points in geometry (table 4.10). This was in addition to the gender gap of about 7 points in geometry that all schools averaged when instructional methods were controlled for (table 4.10).

Besides the one significant variable, several classroom instructional methods variables were almost significant in grades 4 and 12. In grade 4, the more time working on worksheets in math class was almost associated with a larger gender gap in math (table 4.9). In grade 12, taking math tests more often was almost associated with a larger gender gap in 12th grade math (table 4.9), while higher SES schools and schools where the 12th graders worked with objects in math classes, were almost associated with a smaller gender gap in 12th grade geometry (table 4.10).

Even though only one variable was significant in geometry and none in math, controlling for classroom instructional methods did explain 37 percent of the parameter variance in the gender gap in grade 12 math and geometry (tables 4.9 and 4.10). It is possible that the variables in grade 12 that were almost significant contributed to this explanatory power. None of the parameter variance was explained in grades 4 and 8 by this model.

Table 4.9.—Classroom instructional methods predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
GENDER BETA COEFFICIENT			
Intercept	-1.41†	-0.60	-2.52**
Percent African-American	-0.67	-0.84	0.65
Percent Hispanic	-0.33	-1.32	-0.21
Average SES	1.17	-1.58	1.86
Average time spent on:			
Working in small groups	1.23	-0.55	0.96
Working with objects	-0.31	-0.23	1.71
Doing problems on worksheets	-1.73†	-1.38	0.95
Doing problems from textbook	-0.03	-1.20	0.16
Taking math tests	-0.26	-1.07	-2.88†
Using computer	-0.37	0.83	-0.92
Using calculator	0.59	0.67	-0.56
Writing math proofs			-0.71
Formulating own math problems			1.37
OTHER HLM STATISTICS			
Reliability	.15	.21	.12
Parameter variance	23.60*	22.66*	15.26*
Proportion of parameter variance explained	.00	.00	.37

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the gender parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 4.10.—Classroom instructional methods predictors of student-level parameters of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
GENDER BETA COEFFICIENT			
Intercept	-0.01	-1.26	-6.91**
Percent African-American	-0.68	-1.04	1.02
Percent Hispanic	-0.44	-0.83	-0.52
Average SES	1.17	-0.12	3.06†
Average time spent on:			
Working in small groups	1.66	-0.94	1.07
Working with objects	-0.23	-0.03	2.71†
Doing problems on worksheets	-1.18	-0.66	0.90
Doing problems from textbook	-0.48	-1.32	-1.23
Taking math tests	-0.11	-0.27	-1.63
Using computer	-1.17	1.10	-3.21*
Using calculator	0.75	0.79	0.19
Writing math proofs			-0.77
Formulating own math problems			1.95
OTHER HLM STATISTICS			
Reliability	.18	.18	.14
Parameter variance	38.97	28.96	22.33
Proportion of parameter variance explained	.03	.00	.37

[1] All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the gender parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

None of the variables in the three school climate models were significantly associated with the gender gap in grade 12. However, variables in each school climate model were almost significant in predicting the math or geometry gender gap in grade 12. In addition, in two models, controlling for school climate factors explained some parameter variance in the gender gap in geometry in grades 8 and 12.

The following associations were almost significant. In the simple school climate model, schools in which more students enjoyed and felt competent in math tended to have a larger gender gap in 12th grade geometry (table 4.12). However, this was not a robust association, since it disappeared when behavior and safety or academic expectations were added to the model. In the behavior and safety school climate model, schools in which more students experienced disruptions in classes by other students tended to have a smaller gap in math achievement between females and males (table 4.13). In the academic expectations school climate model, schools in which a higher percentage of the 12th grade was in an academic/college prep track tended to have a larger gender gap in math achievement (table 4.15).

These models explained very little parameter variance in the math gender gap in all grades or in the geometry gap in grade 4. However, despite the lack of significant variables, both the behavior and safety model and the academic expectations model explained some part of the parameter variance in the gender gap in geometry in grades 8 and 12. In grade 8, the academic expectations model explained 13 percent of the parameter variance in the geometry gender gap, while in grade 12, the behavior and safety model explained 23 percent of the geometry gender gap.

Table 4.11.—Student climate (math attitudes) predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
GENDER BETA COEFFICIENT			
Intercept	-1.40†	-0.60	-2.64**
Percent African-American	-0.13	-1.30	-0.09
Percent Hispanic	0.05	-1.18	-0.58
Average SES	1.11	-1.52	0.54
Students feel math is useful	1.31	0.12	0.80
Students enjoy and feel competent in math	0.56	-1.02	-1.07
Students disagree that math is more for boys	-0.67	0.28	-0.71
OTHER HLM STATISTICS			
Reliability	.15	.22	.18
Parameter variance	22.88*	24.36*	24.70†
Proportion of parameter variance explained	.01	.00	.00

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the gender parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 4.12.—Student climate (math attitudes) predictors of student-level parameters of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
GENDER BETA COEFFICIENT			
Intercept	-0.03	-1.18	-6.63**
Percent African-American	-0.14	-1.19	0.81
Percent Hispanic	0.00	-0.72	-0.84
Average SES	1.22	-0.13	1.89
Students feel math is useful	1.08	0.43	1.99
Students enjoy and feel competent in math	0.52	-0.62	-2.65†
Students disagree that math is more for boys	-0.33	0.57	-0.84
OTHER HLM STATISTICS			
Reliability	.18	.18	.19
Parameter variance	38.14	29.19	34.05
Proportion of parameter variance explained	.05	.00	.04

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the gender parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 4.13.—Student climate (math attitudes and student behavior and safety) predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
GENDER BETA COEFFICIENT			
Intercept	-1.51†	-0.70	-2.69**
Percent African-American	-0.19	-1.33	-0.24
Percent Hispanic	0.14	-1.00	-0.21
Average SES	1.24	-1.39	0.55
Students feel math is useful	1.14	-0.16	0.77
Students enjoy and feel competent in math	0.72	-0.75	-1.06
Students disagree that math is more for boys	-0.73	0.27	-0.60
Index of problems in the school	0.67		
Percent students enrolled all year	1.45		
Absenteeism in grade		-0.12	-1.78
Students feel classes often disrupted		1.20	2.19†
Students feel unsafe at school		-0.44	-0.88
OTHER HLM STATISTICS			
Reliability	.15	.21	.19
Parameter variance	22.44*	22.53*	25.84†
Proportion of parameter variance explained	.03	.00	.00

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the gender parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education. National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 14.—Student climate (math attitudes and student behavior and safety) predictors of student-level parameters of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
GENDER BETA COEFFICIENT			
Intercept	0.10	-1.30	6.80**
Percent African-American	0.14	-1.34	0.30
Percent Hispanic	0.13	-0.77	-0.65
Average SES	1.15	0.01	2.04
Students feel math is useful	1.10	0.45	1.82
Students enjoy and feel competent in math	0.69	-0.45	-2.60
Students disagree that math is more for boys	-0.35	0.74	-0.73
Index of problems in the school	-0.23		
Percent students enrolled all year	0.87		
Absenteeism in grade		0.51	-1.96
Students feel classes often disrupted		0.40	1.65
Students feel unsafe at school		0.29	0.28
OTHER HLM STATISTICS			
Reliability	.17	.16	.16
Parameter variance	37.07	25.21	27.56
Proportion of parameter variance explained	.07	.05	.23

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the gender parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 4.15.—School climate (math attitudes and academic expectations) predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
GENDER BETA COEFFICIENT			
Intercept	-1.42	-0.47	-2.48*
Percent African-American	-0.50	-1.45	0.84
Percent Hispanic	-0.02	-0.95	-0.32
Average SES	1.33	-0.78	0.58
Students feel math is useful	1.32	-0.18	0.56
Students enjoy and feel competent in math	0.69	-0.90	-0.98
Students disagree that math is more for boys	-0.64	0.12	-0.68
Amount of instruction in math	0.66		
Math identified as a special priority	-1.02		
Percent of 8th grade students taking algebra		-1.30	
Percent of students on academic/college prep			-2.42†
Mean years 12th graders have taken calculus			-1.43
Mean composite math score of grade sample	-0.83	-0.40	2.90
OTHER HLM STATISTICS			
Reliability	.15	.20	.22
Parameter variance	22.03*	21.29*	30.15*
Proportion of parameter variance explained	.04	.05	.00

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the gender parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 4.16.—School climate (math attitudes and academic expectations) predictors of student-level parameters of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
GENDER BETA COEFFICIENT			
Intercept	-0.07	-1.01	-6.54**
Percent African-American	-0.60	-2.19	1.52
Percent Hispanic	-0.01	-0.81	-0.81
Average SES	1.59	1.27	1.26
Students feel math is useful	1.28	0.14	2.08
Students enjoy and feel competent in math	0.56	-0.31	-2.89
Students disagree that math is more for boys	-0.10	0.69	-0.86
Amount of instruction in math	-0.07		
Math identified as a special priority	-1.13		
Percent of 8th grade students taking algebra		-0.43	
Percent of students on academic/college prep			-0.58
Mean years 12th graders have taken calculus			-0.98
Mean composite math score of grade sample	-0.83	-2.44	2.05
OTHER HLM STATISTICS			
Reliability	.18	.15	.21
Parameter variance	37.64	23.11	38.03†
Proportion of parameter variance explained	.06	.13	.00

[1] All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the gender parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

C. Summary

Significant predictors of the gender gap

As summarized in Table 4.17, in most grades and models there were no variables that predicted the variation in the gender gap in math or geometry. This is not surprising given that there was no variance around the gender gap in 12th grade math or in geometry in any grade. However, even where there was parameter variance, as in grades 4 and 8 math, none of the variables in these models explained it. This may be due to fact that the parameter variance in the math gender gap was such a low proportion (15 to 21 percent) of the total variance in these grades. However, there is still a possibility that other variables, or these variables measured in more reliable ways, could contribute some explanatory power. Although the 12th grade gender gap had little significant variation between schools in math or geometry, one variable—the amount of time spend using computers in math class—predicted a larger gender gap in geometry.

Proportion of parameter variance explained and reliability

Tables 4.18 and 4.19 show the proportion of variance explained in the gender gap in math and geometry by the six models. In grades 4 and 8, very little parameter variance was explained. The proportions ranged from 0 to .13. In grade 12, most models also explained none of the parameter variance. However, in the instructional methods model, the parameter variance was not quite significant for 12th grade geometry but the model still explained 37 percent of the variation that was there (table 4.19). In 12th grade math, the parameter variance was significant, and while none of the variables were significant, the model also explained 37 percent of this variation (table 4.18). Thus, the variables in this model, perhaps as a group, have some relationship to the gender gap in 12th grade math and geometry, but without significant variables, it is not possible to tell what that relationship is in math.

As mentioned above, the reliability remained low (15 to 21 percent) in all of the models, and may have contributed to the low amount of parameter variance and proportion of parameter variance explained in most models.

Table 4.17.—Summary of significant school-level variables predicting the gender gap in math and geometry achievement, by HLM model

HLM MODEL Between-school predictors[1]	Grade 4	Grade 8	Grade 12
Significant predictors of the gender gap in math			
AVERAGE GENDER GAP IN MATH	-1.5†	-0.6	-2.5**
SCHOOL RESOURCES MODEL (from detailed model in grade 4)			
School size (number of students)	1.8†		-2.0†
Computers per student	1.6†		
CLASSROOM INSTRUCTIONAL METHODS MODEL			
Average time in math class spent on:			
Doing problems on worksheets	-1.7†		
Taking math tests			-2.9†
SCHOOL CLIMATE (MATH ATTITUDES AND BEHAVIOR AND SAFETY) MODEL			
Students feel classes often disrupted			2.2†
SCHOOL CLIMATE (MATH ATTITUDES AND ACADEMIC EXPECTATIONS) MODEL			
Percent of students on academic/college prep track			-2.4†
Significant predictors of the gender gap in geometry			
AVERAGE GENDER GAP IN GEOMETRY	-0.1	-1.2	-6.5**
CLASSROOM INSTRUCTIONAL METHODS MODEL			
Average SES			3.1†
Average time in math class spent on:			
Working with objects			2.7†
Using computer			-3.2*
SCHOOL CLIMATE (MATH ATTITUDES) MODEL			
Students enjoy and feel competent in math			-2.7†

[1] All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the significant results from the gender parameter equations of the estimated HLM models. For the complete gender equations, see tables 4.4 to 4.16. For the complete HLM models, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 4.18.—Proportion of parameter variance explained by each HLM model for math achievement, grades 4, 8, and 12

Parameter	Student Body Characteristics	School Resources	Classroom Instructional Methods	Math Attitudes (School Climate Models)	Behavior & Safety & Attitudes (School Climate Models)	Academic Expectations & Attitudes (School Climate Models)
	HLM Models					
Grade 4 Mathematics						
GENDER COEFFICIENT	0.03	0.03	0.00	0.01	0.03	0.04
Grade 8 Mathematics						
GENDER COEFFICIENT	0.02	0.00	0.00	0.00	0.00	0.05
Grade 12 Mathematics						
GENDER COEFFICIENT	0.00	0.01	0.37	0.00	0.00	0.00

NOTE: These proportions are calculated from *average* Tau values (averaged across the five plausible values). Negative proportions due to sampling variation have been set to 0.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 4.19.—Proportion of parameter variance explained by each HLM model for geometry achievement, grades 4, 8, and 12

Parameter	Student Body Characteristics	School Resources	Classroom Instructional Methods	Math Attitudes (School Climate Models)	Behavior & Safety & Attitudes (School Climate Models)	Academic Expectations & Attitudes (School Climate Models)
	HLM Models					
Grade 4 Geometry						
GENDER COEFFICIENT	0.07	0.01	0.03	0.05	0.07	0.06
Grade 8 Geometry						
GENDER COEFFICIENT	0.00	0.00	0.00	0.00	0.05	0.13
Grade 12 Geometry						
GENDER COEFFICIENT	0.00	0.06	0.37	0.04	0.23	0.00

NOTE: These proportions are calculated from *average* Tau values (averaged across the five plausible values). Negative proportions due to sampling variation have been set to 0.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Chapter V

School Correlates of Race-Ethnicity Differences in Mathematics Achievement

Overview

This chapter presents the school-level equations that predict the "race-ethnicity gap," or the overall race-ethnicity differences in achievement within schools. The average race-ethnicity gap is expressed in the unconditional equation 2.4, and the models estimated to predict the race-ethnicity gap between schools are expressed in the conditional equation 2.8 (table 5.1). The race-ethnicity equation in the between-school HLM models (equation 2.8) tests the association of various school characteristics with the mathematics and geometry achievement gaps within schools between African-American, Hispanic, and Native American students and white and Asian-American students (table 5.1). This chapter presents the average race-ethnicity gap in mathematics and geometry within schools (tables 5.2 and 5.3) and the results of the six models estimated to predict the race-ethnicity gap in mathematics and geometry (tables 5.4 to 5.16).

Table 5.1.—Equations related to race-ethnicity beta coefficient, or the race-ethnicity gap

Within-school student level equation

$$y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + r_{ij} \qquad (2.1)$$

Between-school school-level equations

Unconditional race-ethnicity equation:

$$\beta_{2j} = \gamma_{20} + u_{2j} \qquad (2.4)$$

Conditional race-ethnicity equation:

$$\beta_{2j} = \gamma_{20} + \gamma_{21}W_{21j} + \gamma_{22}W_{22j} + \ldots + \gamma_{2m}W_{2mj} + u_{2j} \qquad (2.8)$$

NOTE: For explanation of terms in equation, see chapter two.

A. Within-school models

Overall race-ethnicity differences within schools

Tables 5.2 and 5.3 show the results of the unconditional models for math and geometry, respectively. These models provide the average race-ethnicity differences within schools (equation 2.4) for math and geometry. Although this chapter focuse, only on the gender parameter, all the parameters in the models are shown because they function as control variables.

Unlike the gender gap, which barely exists in earlier grades until a gap appears in grade 12, a wide gap exists between African-American, Hispanic, and Native American students and white and Asian-American students in every grade in both math and geometry (tables 5.2 and 5.3). On average, African-American, Hispanic, and Native American students averaged about 14 achievement points below white and Asian-American students in each grade in both math and geometry, controlling for gender, SES, taking algebra in grade 8, and semesters of calculus and geometry, respectively, in grade 12. This difference ranged from about 12.5 achievement points in grade 4 geometry to 15.5 achievement points in grade 8 math, but in all other grades the differences were about 14 points.

Tables 5.2 and 5.3 also include the reliability and the original parameter variance around the race-ethnicity gap in this unconditional model. In grade 4, the reliability was only 7 percent in both math and geometry, so very little variance around the race-ethnicity gap had the potential to be explained. The reliability on race-ethnicity gap in grades 8 and 12 varied between 17 and 26 percent, so less than one quarter or one fifth of the variance around the race-ethnicity gap in these grades was parameter variance and had the potential to be explained.

There was little parameter variance among schools in the race-ethnicity gap in grade 4 math and geometry, although it was almost significantly different from zero. Therefore, the gap of 13 to 15 achievement points between African-American, Hispanic, and Native American students and white and Asian-American students was present in most schools, and a model predicting which types of schools had a larger or smaller race-ethnicity gap might not find enough variation in the gap to make those predictions, especially since the reliability was so low. However, there was wide variation in the size of this gap in grades 8 and 12 math and geometry, as indicated by very significant parameter variances (tables 5.2 and 5.3). Although the parameter variance in grades 8 and 12 was less than one quarter of the total variance, it would be expected that some variables could be found to explain some of that parameter variance.

As with the gender gap, even though there was little variation in the grade 4 race-ethnicity gap, the model predicting this gap was estimated in order to be consistent with the other grades and to illustrate how student-level equations behave when they predict coefficients with little variation. The variation in all grades was modeled by school-level equations that predicted which types of schools had larger or smaller gaps between African-American, Hispanic, and Native American students and white and Asian-American students.

Table 5.2.—Average within-school predictors of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS	(Unit of parameter)	Grade 4	Grade 8	Grade 12
INTERCEPT	(Average Achievement)	211.35**	260.28**	291.98**
GENDER	(1=Female)	-1.46†	-0.63	-2.53**
RACE-ETHNICITY	(1=Afr. Am/Hisp./Nat. Am)	-14.74**	-15.48**	-14.13**
SES	(0=Mean)	6.20**	10.65**	12.36**
TAKING ALGEBRA	(1=If currently taking algebra)		31.66**	
YRS OF CALCULUS	(1=one year)			17.58**
OTHER HLM STATISTICS FOR RACE-ETHNICITY PARAMETER ONLY				
Reliability		.07	.19	.26
Parameter variance		21.41†	44.92**	72.88**

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 5.3.—Average within-school predictors of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS	(Unit of parameter)	Grade 4	Grade 8	Grade 12
INTERCEPT	(Average Achievement)	2213.13**	258.04**	292.08**
GENDER	(1=Female)	-0.14	-1.23	-6.47**
RACE-ETHNICITY	(1=Afr. Am/Hisp./Nat. Am.)	-12.48**	-14.00**	-14.30**
SES	(0=Mean)	4.53**	8.93**	7.49**
TAKING ALGEBRA	(1=If currently taking algebra)		27.84**	
YRS OF GEOMETRY	(1=One year)			23.58**
OTHER HLM STATISTICS FOR RACE-ETHNICITY PARAMETER ONLY				
Reliability		.07	.17	.25
Parameter variance		23.66†	55.30*	90.50**

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

B. Between-school models

Predictors of race-ethnicity differences within schools

Overview

Tables 5.4 to 5.16 show the HLM results for the six models (based on equation 2.8) used to predict the race-ethnicity gap in math and geometry.[63] This section highlights the school-level variables in each model that were found to be significantly related to the race-ethnicity gap. Tables 5.17 and 5.18 summarize the school-level variables that were significantly associated with the race-ethnicity gap in math and geometry achievement and show the number of achievement points in the gap associated with a change in each variable by one standard deviation. Tables 5.19 and 5.20 display the proportion of parameter variance explained for the math and geometry race-ethnicity gap by each model.

Specific Models

The race-ethnicity gap equations yielded variables from the school resources model, the classroom instructional methods model, and one of the school climate models that were significantly associated with the race-ethnicity gap (tables 5.4 to 5.16). However, even models with no significant variables sometimes contributed to explaining the parameter variance, even if there was very little parameter variance.

None of the student body characteristics variables significantly predicted variance in either the math or geometry race-ethnicity gap, and consequently for most grades this model explained little of the parameter variance. However, despite a lack of significant predictors and even a lack of parameter variance in grade 4 math, this model explained 20 percent of the variance in the race-ethnicity gap in 4th grade math and 25 percent of the variance in the race-ethnicity gap in 8th grade math. After explaining this much of the variance, there was still a non-significant amount of variance to explain in grade 4 math, and a significant amount of variance to explain in grade 8 math. In addition, there remained a non-significant amount of variance to explain in grade 4 geometry, and a significant amount of variance to explain in grade 8 geometry and grade 12 math and geometry.

As explained in chapter four, the student body characteristics variables were included in subsequent models despite a lack of significant predictors in all grades and a lack of parameter variance in some grades. These variables were included in the models in all grades to ensure that the theoretical models were consistent, and that the treatment of centered variables was accurate.

[63]The average achievement, gender, and race-ethnicity parameters were random and modeled, while the SES and course-taking parameters were fixed and not modeled. Although all of these parameters were included in the full HLM model, only the equation 2.6, which modeled the race-ethnicity parameter is shown in this chapter. For the full model, see the corresponding HLM tables in Appendix A.

Table 5.4.—Student body characteristics predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.76**	-16.32**	-14.04**
Percent African-American	-0.10	0.13	-0.97
Percent Hispanic	0.53	1.78	0.35
Average SES	0.79	-1.67	-0.08
OTHER HLM STATISTICS			
Reliability	.06	.17	.26
Parameter variance	15.97†	35.96**	74.72**
Proportion of parameter variance explained	.25	.20	.00

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the race-ethnicity parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 5.5—Student body characteristics predictors of student-level parameters of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-12.84**	-14.54**	-14.15**
Percent African-American	-0.53	-1.30	-2.00
Percent Hispanic	0.57	1.58	0.54
Average SES	0.56	-2.07	-2.80
OTHER HLM STATISTICS			
Reliability	.12	.16	.23
Parameter variance	42.99†	51.58†	88.02**
Proportion of parameter variance explained	.00	.07	.03

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the race-ethnicity parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

In the school resources model, several variables were associated with the race-ethnicity gap in grade 12 math, but none were associated with the race-ethnicity gap in math in grades 4 or 8 or in geometry in any grades (tables 5.6 to 5.8). Consequently, in grades 4 and 8, this model explained little parameter variance in the race-ethnicity gap in either math or geometry. However, in grade 12 math, reflecting the presence of significant predictors, this model explained 30 percent of the variance (table 5.6). And in grade 12 geometry, despite no significant predictors, this model explained 18 percent of the variance (table 5.7).

In grade 12 math, a higher student/teacher ratio and more instructional funds per student in a district were associated with a smaller gap between African Americans/Hispanics/Native Americans and whites/Asian-Americans (table 5.6). The gap of about 14 points between African-American, Hispanic, and Native American students and white and Asian-American students in 12th grade math achievement was smaller by about 5 points for every standard deviation above the average student/teacher ratio, and the gap was smaller by about 4 points for every standard deviation above the average level of funds per student in the district (table 5.6). One other variable was almost significant in this model for grade 12 math—a larger school size in number of students was almost significantly associated with a larger race-ethnicity gap in math (table 5.6).

Table 5.6.—School resources predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.69**	-16.26**	-14.06**
Percent African-American	0.45	0.27	-0.50
Percent Hispanic	0.63	1.61	-0.18
Average SES	0.99	-1.70	0.30
School size (number of students)	-0.61	-0.56	-3.18†
Student/teacher ratio	0.96	0.81	4.57*
District instructional funds/student	-1.34	0.73	4.17*
OTHER HLM STATISTICS			
Reliability	.08	.17	.19
Parameter variance	21.82*	39.18**	50.67**
Proportion of parameter variance explained	.00	.13	.30

[1] All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the race-ethnicity parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 5.7.—School resources predictors of student-level parameters of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-12.97**	-14.26**	-13.63**
Percent African-American	0.01	-1.09	-1.59
Percent Hispanic	0.43	1.48	0.43
Average SES	0.92	-2.06	-2.56
School size (number of students)	0.02	-1.03	-3.07
Student/teacher ratio	1.31	0.11	3.03
District instructional funds/student	-2.05	1.33	3.11
OTHER HLM STATISTICS			
Reliability	.13	.15	.21
Parameter variance	46.09†	50.16†	74.51**
Proportion of parameter variance explained	.00	.09	.18

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the race-ethnicity parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 5.8.—Detailed school resources predictors of student-level parameters of grade 4 math and geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Math Grade 4	Geometry Grade 4
RACE-ETHNICITY BETA COEFFICIENT		
Intercept	-14.75**	-12.58**
Percent African-American	0.18	-0.44
Percent Hispanic	0.66	0.56
Average SES	1.25	1.35
School size (number of students)	-0.02	0.38
Student/teacher ratio	1.28	1.62
District instructional funds/student	-1.07	-1.71
Computers per student	1.53	1.22
Percent use computers as part of math instruction	-1.52	-2.58
OTHER HLM STATISTICS		
Reliability	.06	.12
Parameter variance	16.06†	42.32
Proportion of parameter variance explained	.25	.00

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the race-ethnicity parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Several variables in the classroom instructional methods model in grades 4, 8, and 12 were associated with the race-ethnicity gap in math and geometry achievement. In grade 4, schools in which 4th graders spent a higher than average time on worksheets in math class had a lower race-ethnicity gap in math and geometry achievement than other schools (tables 5.9 and 5.10). For every standard deviation above the average time on worksheets in grade 4 math classes, the average race-ethnicity gap of about 15 points in math achievement was reduced by an average of about 3 points, and the race-ethnicity gap of about 13 points in geometry achievement was reduced by an average of about 3 points.

In grade 8, schools in which 8th graders spent a higher than average time working with objects (rulers, blocks, and solids) in math class had a smaller race-ethnicity gap in math achievement than other schools (table 5.9). This variable reduced an average gap of 15.5 points between African-American, Hispanic, and Native American students and white/Asian-American students in math achievement by about 3 points. In geometry, this same variable was almost significant by a similar amount.

In grade 12, no instructional methods were significantly related to the race-ethnicity gap in either math or geometry. However, in geometry, there was one association that was almost significant. Schools in which 12th graders spent more than the average time in math class taking math tests averaged a larger race-ethnicity gap in geometry achievement (table 5.10).

The proportion of parameter variance explained by this model in some ways reflected the grades that had significant predictors, but in some ways did not. For grades 4 and 8 math, where there was one significant predictor each, this model explained about one third of the variance (table 5.9). However, in grade 4 geometry, which also had a significant predictor, no variance was explained by this model (table 5.10). In grade 8 geometry, where one predictor was almost significant, 15 percent of the variance was explained, but in grade 12 geometry, where one predictor was almost significant, only 4 percent of the variance was explained (table 5.10). These differences in the proportion of variance explained did not reflect differences in the reliability in these grades.

Table 5.9.—Classroom instructional methods predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.69**	-16.00**	-14.88**
Percent African-American	0.56	1.19	0.23
Percent Hispanic	1.26	1.49	0.78
Average SES	0.84	-1.81	2.94
Average time spent on:			
Working in small groups	0.47	2.53	0.58
Working with objects	-0.86	3.16*	1.04
Doing problems on worksheets	2.63*	-0.15	2.05
Doing problems from textbook	1.00	-1.37	-4.36
Taking math tests	-0.52	-1.10	-2.65
Using computer	-0.97	-0.74	-3.77
Using calculator	-0.54	-1.27	-2.41
Writing math proofs			-0.20
Formulating own math problems			1.53
OTHER HLM STATISTICS			
Reliability	.06	.15	.26
Parameter variance	14.37†	34.12**	75.43**
Proportion of parameter variance explained	.33	.30	.00

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the race-ethnicity parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 5.10.—Classroom instructional methods predictors of student-level parameters of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-12.60**	-14.01**	-13.99**
Percent African-American	0.11	0.12	0.64
Percent Hispanic	1.42	1.29	0.83
Average SES	0.88	-2.09	-1.29
Average time spent on:			
Working in small groups	-0.30	2.36	-0.52
Working with objects	-0.64	3.24†	-0.15
Doing problems on worksheets	3.03*	-0.93	3.17
Doing problems from textbook	0.67	-0.84	-2.06
Taking math tests	-0.21	-1.54	-6.71†
Using computer	-1.57	-0.01	0.95
Using calculator	-2.13	-1.43	-0.16
Writing math proofs			0.93
Formulating own math problems			-0.05
OTHER HLM STATISTICS			
Reliability	.11	.15	.24
Parameter variance	40.14	46.93	86.97**
Proportion of parameter variance explained	.00	.15	.04

[1] All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the race-ethnicity parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Finally, in the school climate models, none of the school climate variables were significantly associated with the race-ethnicity gap. However, in grade 8 math, the percentage of Hispanic students in a school became significantly associated with a smaller race-ethnicity gap in the third school climate model (math attitudes plus academic expectations), after being almost significant in the first school climate model (math attitudes only) (tables 5.11 and 5.15). Controlling for attitudes towards math and for the percentage of 8th graders taking algebra, schools with one standard deviation above the average percentage of Hispanics averaged about a 3 point smaller race-ethnicity gap than the average race-ethnicity gap of 15.5 points (table 5.15). However, this variable was not significant in the second school climate model (math attitudes plus student safety and behavior). Since percentage of Hispanic students had never been a significant factor in any other model, its significance here casts doubt on the importance of this finding, and illustrates the sensitivity of the significance of coefficients to the other variables in the model.

Several variables in each grade were almost significantly associated with the race-ethnicity gap in math and geometry achievement. For grade 4 math, in the math attitudes and academic expectations model, the percentage of students who felt that math is useful was almost associated with a wider race-ethnicity gap in math achievement (table 5.15). In addition, higher than average amounts of instruction in math in grade 4 were almost associated with smaller gaps in math and geometry achievement, respectively (tables 5.15 and 5.16).

For grade 8 math, in the student behavior and academic expectations models, associations that were almost significant were the following. The more 8th graders who enjoyed and felt competent in math, the smaller the race-ethnicity gap in math achievement (table 5.15). However, in schools where more 8th graders were taking algebra, the math race-ethnicity gap was larger (table 5.15).

In grade 12, two variables were almost significant. In all three of the school climate models for math and in the academic expectations model for geometry, schools in which more 12th graders hold positive attitudes towards females and math averaged a larger race-ethnicity gap in math and geometry (tables 5.11, 5.13, 5.15, and 5.16). In addition, in the student behavior and safety model, schools with higher average SES levels averaged higher race-ethnicity gaps in geometry (table 5.14).

The proportion of variance explained by each school climate model bore no relation to the number of significant or almost-significant variables in the model or to the reliability in the unconditional model. The academic expectations model for grade 4 math explained the most variance—50 percent—although grade 4 had only 2 almost-significant variables and a reliability of only .04 in that model. However, half of that percent, or 25 percent of the variance had been explained by the student body characteristics variables. For grade 8 math, these three models explained about 25 percent of the variance in the race-ethnicity gap. Although only one of the models had a significant variable, there was at least one almost-significant variable in each model. Still, most of that percent, or 20 percent of the variance had been explained by the student body characteristics variables. For grade 12, which had almost-significant variables in each model, these models explained none of the variance in the math race-ethnicity gap, explained little variance in the geometry race-ethnicity gap with the first and third models, but explained 19 percent of the variance in the geometry race-ethnicity gap with the second model.

Table 5.11.—Student climate (math attitudes) predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.76**	-16.27**	-14.17**
Percent African-American	-0.25	0.07	-1.53
Percent Hispanic	0.62	1.98†	0.40
Average SES	1.39	-1.50	-0.48
Students feel math is useful	-1.70	0.32	-1.31
Students enjoy and feel competent in math	-0.87	2.81	0.00
Students disagree that math is more for boys	0.10	1.25	-3.13†
OTHER HLM STATISTICS			
Reliability	.08	.14	.28
Parameter variance	19.95†	31.52*	83.95**
Proportion of parameter variance explained	.07	.24	.00

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the race-ethnicity parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 5.12.—Student climate (math attitudes) predictors of student-level parameters of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-12.81**	-14.54**	-14.04**
Percent African-American	-0.74	-1.45	-1.93
Percent Hispanic	0.59	1.92	0.66
Average SES	1.09	-1.84	-3.59
Students feel math is useful	-1.25	-0.36	-1.87
Students enjoy and feel competent in math	-0.91	4.10	-2.83
Students disagree that math is more for boys	-0.05	1.25	-3.07
OTHER HLM STATISTICS			
Reliability	.13	.14	.23
Parameter variance	47.50	44.16†	83.02**
Proportion of parameter variance explained	.00	.20	.08

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the race-ethnicity parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 5.13.—Student climate (math attitudes and student behavior and safety) predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.93**	-16.75**	-12.80**
Percent African-American	-0.20	-0.52	-0.80
Percent Hispanic	0.68	1.71	0.37
Average SES	1.17	-1.18	-1.11
Students feel math is useful	-1.81	0.26	-0.93
Students enjoy and feel competent in math	-0.89	3.25†	-0.96
Students disagree that math is more for boys	-0.03	1.80	-3.37†
Index of problems in the school	-0.83		
Percent students enrolled all year	0.19		
Absenteeism in grade		1.40	-0.05
Students feel classes often disrupted		0.95	-2.90
Students feel unsafe at school		1.35	-1.56
OTHER HLM STATISTICS			
Reliability	.07	.15	.26
Parameter variance	18.32†	33.14**	72.87**
Proportion of parameter variance explained	.14	.26	.00

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the race-ethnicity parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 5.14.—Student climate (math attitudes and student behavior and safety) predictors of student-level parameters of geometry achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-12.61**	-14.96**	-12.92**
Percent African-American	-0.47	-1.79	-1.72
Percent Hispanic	0.64	1.64	0.88
Average SES	0.66	-1.71	-4.45†
Students feel math is useful	-1.22	-0.18	-1.54
Students enjoy and feel competent in math	-0.83	4.33	-3.63
Students disagree that math is more for boys	-0.03	1.61	-2.98
Index of problems in the school	-1.57		
Percent students enrolled all year	-0.60		
Absenteeism in grade		1.43	2.78
Students feel classes often disrupted		0.19	-1.59
Students feel unsafe at school		1.00	-1.45
OTHER HLM STATISTICS			
Reliability	.13	.15	.21
Parameter variance	49.86†	49.06	73.51**
Proportion of parameter variance explained	.00	.11	.19

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the race-ethnicity parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 5.15.—School climate (math attitudes and academic expectations) predictors of student-level parameters of math achievement, grades 4, 8, and 12

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-15.00**	-15.77**	-13.98**
Percent African-American	-0.60	-0.61	0.89
Percent Hispanic	0.53	2.55*	0.18
Average SES	1.28	0.69	-2.93
Students feel math is useful	-1.98†	-0.54	-0.90
Students enjoy and feel competent in math	-0.78	3.35†	-1.44
Students disagree that math is more for boys	0.13	1.00	-3.63†
Amount of instruction in math	1.88†		
Math identified as a special priority	-0.32		
Percent of 8th grade students taking algebra		-3.02†	
Percent of students on academic/college prep			-3.36
Mean years 12th graders have taken calculus			0.26
Mean composite math score of grade sample	-0.36	-1.47	6.16
OTHER HLM STATISTICS			
Reliability	.04	.15	.29
Parameter variance	10.77†	33.83**	85.40**
Proportion of parameter variance explained	.50	.25	.00

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the race-ethnicity parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 5.16.—School climate (math attitudes and academic expectations) predictors of student-level parameters of geometry achievement, grades 4, 8, and 12—Continued

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 4	Grade 8	Grade 12
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-12.73**	-14.03**	-13.76**
Percent African-American	-1.09	-3.31	0.35
Percent Hispanic	0.46	1.97	1.02
Average SES	1.04	1.08	-4.03
Students feel math is useful	-1.45	-1.06	-2.48
Students enjoy and feel competent in math	-0.50	4.98	-3.07
Students disagree that math is more for boys	0.17	1.61	-3.44†
Amount of instruction in math	2.60†		
Math identified as a special priority	-0.37		
Percent of 8th grade students taking algebra		-1.26	
Percent of students on academic/college prep			-5.34
Mean years 12th graders have taken calculus			-2.12
Mean composite math score of grade sample	-0.78	-4.52	6.58
OTHER HLM STATISTICS			
Reliability	.11	.17	.24
Parameter variance	37.84	57.89†	86.81*
Proportion of parameter variance explained	.00	.00	.04

[1]All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the race-ethnicity parameter equation of the estimated HLM model. For the complete HLM model, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

C. Summary

Significant predictors of the race-ethnicity gap

As summarized in Tables 5.17 and 5.18, only four variables were found to predict the race-ethnicity gap in math, and only one variable significantly predicted the race-ethnicity gap in geometry, although the gap was substantial in both areas. The types of schools that had lower race-ethnicity gaps in grade 4 math and geometry were those in which students did problems in worksheets more often. In grade 8, the gap in math was lower in schools in which students worked with objects more often and had a high percentage of Hispanic students, although this last association was only true in one out of the six models tested, so it is not considered robust. In grade 12, the race-ethnicity gap was smaller in schools in which there was a higher student/teacher ratio and more district instructional funds per student.

Proportion of parameter variance explained and reliability

Tables 5.19 and 5.20 show the proportion of variance explained in the race-ethnicity gap in math and geometry by the six models. This proportion ranged from 0 to 50 percent for math achievement and 0 to 20 percent for geometry achievement, and it had little relation to the number of significant predictors or level of reliability of the models.

In grade 4, the reliability, or proportion of the total variance that was parameter variance, was only .07 in the race-ethnicity gap in both math and geometry. Yet the student body characteristics model explained 25 percent and the classroom instructional methods model explained 33 percent of that variance in the race-ethnicity gap in math, although only the classroom instructional methods model had a significant predictor. Although the instructional methods model also had a significant predictor for the geometry race-ethnicity gap, this model explained none of the variance in that area. In contrast, the academic expectations model in grade 4 explained 50 percent of the variance in the race-ethnicity gap in math, although this model had no significant predictors.

In grade 8, the reliability was .19 in the math race-ethnicity gap and .17 in the geometry race-ethnicity gap. The classroom instructional methods model explained 30 percent of the variance in math, and there was one significant predictor in this model. The three school climate models explained about 25 percent of the variance in math, but none of the school climate variables were significant.

In grade 12, the reliability was .26 in the math race-ethnicity gap and .25 in the geometry race-ethnicity gap. Only the school resources model explained any variance in math, and it explained 30 percent of the variance. This model contained the two variables that were significant predictors for the grade 12 math race-ethnicity gap. For the race-ethnicity gap in geometry, school climate models explained the most variance in grades 8 and 12, but they still explained only 19–20 percent.

Table 5.17.—Summary of significant school-level variables predicting the race-ethnicity gap in math achievement, by HLM model

HLM MODEL Between-school predictors[1]	Grade 4	Grade 8	Grade 12
AVERAGE RACE-ETHNICITY GAP IN MATH	-14.7**	-15.5**	-14.1**
SCHOOL RESOURCES MODEL			
School size (number of students)			-3.2†
Student/teacher ratio			4.6*
District instructional funds/student			4.2*
CLASSROOM INSTRUCTIONAL METHODS MODEL			
Average time in math class spent on:			
Working with objects		3.2*	
Doing problems on worksheets	2.6*		
SCHOOL CLIMATE (MATH ATTITUDES) MODEL			
Percent Hispanic		2.0†	
Students disagree that math is more for boys			-3.1†
SCHOOL CLIMATE (MATH ATTITUDES AND BEHAVIOR AND SAFETY) MODEL			
Students enjoy and feel competent in math		3.3†	
Students disagree that math is more for boys			-3.4†
SCHOOL CLIMATE (MATH ATTITUDES AND ACADEMIC EXPECTATIONS) MODEL			
Percent Hispanic		2.6*	
Students feel math is useful	-2.0†		
Students enjoy and feel competent in math		3.4†	
Students disagree that math is more for boys			-3.6†
Amount of instruction in math	1.9†		
Percent of 8th grade students taking algebra		-3.0†	

[1] All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the significant results from the race-ethnicity parameter equations of the estimated HLM models. For the complete race-ethnicity equations, see tables 5.4 to 5.16. For the complete HLM models, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 5.18.—Summary of significant school-level variables predicting the race-ethnicity gap in geometry achievement, by HLM model

HLM MODEL Between-school predictors[1]	Grade 4	Grade 8	Grade 12
AVERAGE RACE-ETHNICITY GAP IN GEOM.	-12.5**	-14.0**	-14.3**
CLASSROOM INSTRUCTIONAL METHODS MODEL			
Average time in math class spent on:			
Working with objects		3.2†	
Doing problems on worksheets	3.0*		
Taking math tests			-6.7†
SCHOOL CLIMATE (MATH ATTITUDES AND BEHAVIOR AND SAFETY) MODEL			
Average SES			-4.5†
SCHOOL CLIMATE (MATH ATTITUDES AND ACADEMIC EXPECTATIONS) MODEL			
Students disagree that math is more for boys			-3.4†
Amount of instruction in math	2.6†		

[1] All between-school independent variables have been standardized. See chapter two for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the significant results from the race-ethnicity parameter equations of the estimated HLM models. For the complete race-ethnicity equations, see tables 5.4 to 5.16. For the complete HLM models, see the corresponding tables in Appendix A.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 5.19.—Proportion of parameter variance explained by each HLM model for math achievement, grades 4, 8, and 12

Parameter	HLM Models					
	Student Body	School Resources	Instructional Methods	Math Attitudes	Behavior & Safety	Academic Expect.
Grade 4 Mathematics						
RACE-ETHNICITY COEFF.	0.25	0.00	0.33	0.07	0.14	0.50
Grade 8 Mathematics						
RACE-ETHNICITY COEFF.	0.20	0.13	0.30	0.24	0.26	0.25
Grade 12 Mathematics						
RACE-ETHNICITY COEFF.	0.00	0.30	0.00	0.00	0.00	0.00

NOTE: These proportions are calculated from *average* Tau values (averaged across the five plausible values). Negative proportions due to sampling variation have been set to 0.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Table 5.20.—Proportion of parameter variance explained by each HLM model for geometry achievement, grades 4, 8, and 12

Parameter	HLM Models					
	Student Body	School Resources	Instructional Methods	Math Attitudes	Behavior & Safety	Academic Expect.
Grade 4 Geometry						
RACE-ETHNICITY COEFF.	0.00	0.00	0.00	0.00	0.00	0.00
Grade 8 Geometry						
RACE-ETHNICITY COEFF.	0.07	0.09	0.15	0.20	0.11	0.00
Grade 12 Geometry						
RACE-ETHNICITY COEFF.	0.03	0.18	0.04	0.08	0.19	0.04

NOTE: These proportions are calculated from *average* Tau values (averaged across the five plausible values). Negative proportions due to sampling variation have been set to 0.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

Chapter VI

Exploratory Analyses of the Gender and Race-Ethnicity Gaps

A. Overview

Two exploratory analyses were performed that involved altering the models predicting the gender and race-ethnicity gaps. The first analysis used the deviance and Chi-square statistics to determine whether the student-level race-ethnicity coefficient could be fixed, and then examined the effect of fixing it. The second analysis took the course-taking control variables out of the student-level equations in grade 8 and 12 and examined the effects of this change. The results of these changes were measured on the predictors and parameter variance (Tau) of the gender and race-ethnicity gaps. While both analyses were fairly simple variations of the original models, the execution of the first analysis was greatly complicated by the use of the NAEP plausible values.

B. The effect of fixing the race-ethnicity coefficient

The deviance statistic and the Chi-square statistic

The deviance statistic is the -2 (log likelihood). It can be used to determine whether there is any significant parameter variance in random models by comparing it to the deviance statistic in a model where that parameter is fixed. This method is similar to the use of the *F* statistic to compare OLS equations to determine whether a fuller model has added significantly to explaining the variance or not. If it does not add anything, then the reduced model may just as well be used. The deviance statistic allows comparisons of models with a parameter fixed and its Tau at zero versus models with the same parameter random and its Tau at non-zero. Unlike the *F* statistic, the deviance statistic has no significance in itself, but the *difference* between the deviance statistics from each model is either significantly different from zero or not. The deviance statistics are compared using a likelihood ratio test.

A significant difference between deviance statistics indicates that one ought to keep the parameter random and Tau non-zero. If it is not a significant difference, that is, if the parameter variance is not significantly different from 0, then the parameter may just as well be fixed. Fixing a parameter is efficient computationally because then the variable is not included in the variance/covariance matrix. In addition, fixed control variables can be added without losing degrees of freedom in the within-school equation. Fixing and not modeling it allows the school-level models to predict only the parameters with variance. Then these models are not affected by the parameter with no variance or by its predictors.

However, other options include fixing the parameter but modeling it anyway, which will explain the nonrandom but varying variation of the parameter. In this case, the parameter is not included in the variance/covariance matrix, but its own variation is modeled. Another option is to keep the parameter random and model whatever variation there is.

The deviance statistic and the Chi-square statistic use different information to test for parameter variance significance. The deviance statistic uses the variance/covariance matrix from the entire model and from all the schools, while the Chi-square statistic uses only the variance from that parameter, and thus can only use information from schools with variance on that parameter (see below). For this reason, researchers are encouraged to use the deviance statistic rather than the Chi-square statistic as a final test for determining the parameter variance significance. Unless one is completely theory-driven, one should only fix the parameter if the deviance statistic shows no difference. However, this guideline can often produce the following outcomes:

- Fixing variables of interest and modeling their nonrandom variation, in some cases even if a Chi-square test shows that there is parameter variance

- Not fixing a variable even if a Chi-square test shows there is no parameter variance

- Keeping random some of the control variables but not modeling them

These outcomes can result in a very mixed, confusing set of models. The best procedure to take is to decide on the variance-covariance of a parameter (fixed versus random) based on strong theoretical grounds, rather than on exploratory results alone. If there is no parameter variance, one can either fix it or keep it random, and still either model or not model the nonrandom variation. However, since this method is inefficient, one would not want to include too many of these nonfixed, nonrandom parameters in the equation.

The advantage of the deviance statistic with segregated schools (all white or all minority)

HLM uses the variance/covariance information from all the schools to produce the Gamma values and statistics, which predict the intercept and each Beta coefficient of the student-level equations. One of these Beta coefficients is the race-ethnicity gap. If only the OLS equations were used to estimate this within-school Beta coefficient, then these equations would only be available for schools with both minorities and whites, rather than segregated (all-white or all-minority) schools. Segregated schools would not be included because the race-ethnicity variable would be a constant of either white or minority in those schools. Thus, there are no OLS equations for segregated schools because, like any OLS model with an independent variable with no variation, the equation cannot be calculated.

However, in HLM, variance/covariance information from the segregated schools is still included in the average race-ethnicity gap coefficient due to the mixed model algorithm. HLM calculates the Gammas, the Taus, and the deviance statistics using all the schools in the following way. It performs what is called "borrowing strength" from the schools and equations that are available for the schools and equations that are not. Therefore, while data from segregated schools would add no information about the race-ethnicity gap, such data could add information about gender, SES, and average achievement. Thus, HLM uses this information and calculates values based on what is known from the other equations.

Therefore, Tau is calculated using all schools, but the Chi-square test of Tau, the reliability values, and the starting values are calculated only for those schools with OLS equations. The portion of Tau that is evaluated with the Chi-square test comes only from the diverse schools. This explains the warning statement under the Chi-square table in the HLM/2L program, which reminds analysts not to rely on the Chi-square test since it is based only on schools with OLS equations. The same warning is printed under the OLS equations and the reliability values. Thus, the major reason for using the deviance statistic is that it evaluates the parameter variance (Tau) using information from all the schools.

Testing the parameter variance

In the original within-school equations, it was decided a priori to keep gender and race-ethnicity random and to fix SES in grade 4 and to fix SES and coursetaking in grades 8 and 12. However, these decisions did not necessarily reflect the actual amount of parameter variance around each coefficient. Another way to decide which variables to fix is to use the Chi-square statistic and the deviance statistic to test whether there is any variance in Tau to model.

In this study, the average Chi-square estimate of the parameters indicated that in grade 4, the variance of race-ethnicity did not vary, indicating that race-ethnicity could be fixed. The deviance test was performed in order to confirm this. A grade 4 model was run with all variables random and compared with the same model with race-ethnicity fixed. The goal of this exploratory analysis was to show how different the deviance vs. the Chi-square test was for determining the significance of the Tau, and to determine whether the Tau of the gender coefficient changed when race-ethnicity was fixed. The following models were tested:

Model I. Grade 4 Unconditional Model: Parameter variance was random for all parameters.

$$y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + r_{ij} \tag{6.1}$$

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad \text{(Intercept equation)} \tag{6.2}$$

$$\beta_{1j} = \gamma_{10} + u_{1j} \quad \text{(Gender gap equation)} \tag{6.3}$$

$$\beta_{2j} = \gamma_{20} + u_{2j} \quad \text{(Race-ethnicity gap equation)} \tag{6.4}$$

$$\beta_{3j} = \gamma_{30} + u_{3j} \quad \text{(SES equation)} \tag{6.5}$$

where:
i represents the i^{th} student
j represents the j^{th} school
y_{ij} represents the achievement score of the i^{th} student in the j^{th} school
β_{0j} is the intercept, or the average achievement in the j^{th} school
β_{1j} is the Beta coefficient for gender in the j^{th} school
β_{2j} is the Beta coefficient for race-ethnicity in the j^{th} school
β_{3j} is the Beta coefficient for SES in the j^{th} school
X_{1ij} represents the gender of the i^{th} student in the j^{th} school
X_{2ij} represents race-ethnicity of the i^{th} student in the j^{th} school
X_{3ij} represents SES of the i^{th} student in the j^{th} school
r_{ij} is random error in the j^{th} school.
p is the number of within-school parameter equations (from 0 to 3 in this example)
γ_{p0} is the intercept, or the average within-school parameter value in the p^{th} equation
u_{pj} is random error in the p^{th} equation.

Model I was compared to the following model:

Model II. Grade 4 Unconditional Model: Parameter variance was fixed for race-ethnicity.

$$y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \beta_{3j}X_{3ij} + r_{ij} \quad (6.1)$$

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad \text{(Intercept equation)} \quad (6.2)$$

$$\beta_{1j} = \gamma_{10} + u_{1j} \quad \text{(Gender gap equation)} \quad (6.3)$$

$$\beta_{2j} = \gamma_{20} \quad \text{(Race-ethnicity gap equation)} \quad (6.4)$$

$$\beta_{3j} = \gamma_{30} + u_{3j} \quad \text{(SES equation)} \quad (6.5)$$

Tables 6.1 and 6.2 show the parameter variance, the Chi-square test and its probability value, and the deviance statistics for Model I and Model II. In addition, the Chi-square test of the difference between the deviance statistics for each model is shown in table 6.2. These tables illustrate how the Chi-square test on the Tau can differ from the Chi-square test on the deviance statistic for each plausible value.[64]

[64]This exploratory analysis was performed on three of the five available plausible values because two of the values would not produce estimates for the unconditional model with all parameters random. We followed the same procedure as for five plausible values, which was to average the statistics of the three plausible values. However, we also show the results for each of the three plausible values.

Table 6.1.—Parameter variance (Tau), Chi-square test for Tau > 0, probability value, and deviance statistic, by plausible values and average: Unconditional model predicting math composite with no parameter variances fixed, grade 4 (Model I)

MODEL PARAMETER STATISTIC	Math Composite Plausible Value Used			
	2	3	5	Average
GENDER BETA COEFFICIENT				
Tau	24.5	1.9	19.70	15.4
Chi-square	256.6	248.9	258.7	254.7
Probability Tau > 0	.034	.067	.028	.043
RACE-ETHNICITY COEFFICIENT				
Tau	7.5	42.5	20.8	23.6
Chi-square	266.6	256.5	229.6	250.9
Probability Tau > 0	.012	.034	.266	.104
SES COEFFICIENT				
Tau	20.4	21.0	16.6	19.4
Chi-square	263.1	251.5	235.0	249.9
Probability Tau > 0	.018	.054	.191	.088
DEVIANCE STATISTIC FOR EQUATION WITH NO PARAMETERS FIXED				
Deviance1	43841.3	43986.8	43913.8	
Degrees of freedom	11	11	11	

NOTE: Chi-square test of Tau > 0 based on 218 out of 256 schools.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

In table 6.1, the Chi-square test of the Tau indicates that for two out of three plausible values, there was significant variation in Tau for the race-ethnicity coefficient, so race-ethnicity should not be fixed. However, for the other plausible value, there was no variation on that coefficient, and these three Chi-squares averaged to a Chi-square that was not significant. When race-ethnicity was fixed and a Chi-square test was used to compare the deviance statistic of that equation to the deviance statistic of the model when race-ethnicity was not fixed, this test indicated that there was no significant difference in the Tau (table 6.2). Therefore, race-ethnicity could be fixed if efficiency was wanted. It is interesting to note that when race-ethnicity was fixed, the Taus of gender and SES, based on the Chi-square test of Tau > 0, were somewhat larger and yet more significant. This may be the result of the inclusion of almost all of the schools (255 out of 256 schools) since by fixing race-ethnicity, segregated schools were included. Thus, fixing race-ethnicity may have produced a more accurate Chi-square test of the Tau significance for the other parameters.

Table 6.2.—Parameter variance (Tau), Chi-square test for Tau > 0 and probability value, deviance statistic, and Chi-square test for difference between deviance statistics>0 and probability value, by plausible values and average of statistics: Unconditional model predicting math composite with race-ethnicity fixed, grade 4 (Model II)

MODEL PARAMETER STATISTIC	Math Composite Plausible Value Used 2	3	5	Average
GENDER BETA COEFFICIENT				
Tau	27.7	23.6	26.6	26.0
Chi-square	316.8	312.6	319.3	316.2
Probability Tau > 0	.005	.008	.004	.006
SES COEFFICIENT				
Tau	23.0	15.9	14.6	17.8
Chi-square	316.0	305.2	292.9	304.7
Probability Tau > 0	.006	.017	.051	.025
DEVIANCE STATISTIC FOR EQUATION WITH RACE-ETHNICITY FIXED				
Deviance2	43845.8	43986.3	43917.2	
Degrees of freedom	7	7	7	
TEST FOR DIFFERENCE BETWEEN DEVIANCE1 AND DEVIANCE2				
Chi-square	4.46	.51	3.35	2.77
Degrees of freedom	4	4	4	4
Probability	.35	>.5	>.5	>.5

NOTE: Chi-square test of Tau > 0 now based on 256 schools. Chi-square test comparing deviance statistics based on 256 schools.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade, 8th Grade, and 12th Grade Students, Restricted-Use Data Base.

This same exploratory analysis was performed on the classroom instructional methods model to determine whether the deviance test continued to indicate that the race-ethnicity coefficient should be fixed, and to confirm that the estimates for the gender equation do not change when race-ethnicity is fixed but modeled and SES is not fixed. This analysis (not shown) produced similar results as the unconditional model. The Chi-square test of Tau > 0 indicated that for two out of three plausible values, race-ethnicity could be fixed, and the deviance statistics test confirmed this for all three plausible values. As in the unconditional model, when the Tau of the race-ethnicity coefficient was fixed, the Tau of the gender coefficient varied more than when race-ethnicity was random and the segregated schools were not included.

As expected, the Gamma estimates of the gender and race-ethnicity equations did not change when the race-ethnicity Tau was fixed, because Gamma estimates are based on all of the schools. Gamma estimates depend only on the other school-level variables and the student-level models in each school-level equation. Therefore, the Gamma estimates were not affected by fixing or not fixing race-ethnicity or by fixing or not fixing SES.

C. Taking out the controls for coursetaking in grades 8 and 12

Coursetaking was included as a control variable in the grade 8 and 12 student-level equations because it is known to be associated with gender differences in math achievement. By controlling for it, it was hoped that the effects of gender separate from coursetaking could be estimated. However, although average gender differences in achievement were found when coursetaking was controlled for in grade 12, no average gender differences were found in grade 8, and only one school-level variable in grade 12 could explain the variation in gender differences in either grade. Since the coursetaking coefficient was always large and significant, this variable may have been a proxy for much of the difference in achievement by gender, and since it was not modeled, the school correlates of these differences were hidden.

These results suggested that an exploratory model without coursetaking in the within-school equation might produce different results for gender than when coursetaking was in the model. Estimating gender differences without controlling for coursetaking differences would test for gender differences that would include coursetaking differences by gender, and this model would be expected to exacerbate gender differences.

To test this hypothesis, coursetaking was taken out of the grade 8 and grade 12 overall math and grade 12 geometry within-school equation. SES was still fixed at the within-school level. The between-school model of classroom instructional methods was estimated. This model was chosen because it contained the most variables in any model that were hypothesized to be related to the gender coefficient.

Table 6.3 shows the results of the between-school model for the gender and race-ethnicity coefficients. Also shown in parentheses are the coefficients from the gender and race equation in the original classroom instructional methods model, which includes coursetaking (tables 4.9, 4.10, 5.9, and 5.10). The results in table 6.3 are compared to the results from this previous model.

Table 6.3.—Classroom instructional methods predictors of student-level parameters of grade 8 math, grade 12 math, and grade 12 geometry, after coursetaking has been removed from the within-school equation

WITHIN-SCHOOL PARAMETERS Between-school predictors[1]	Grade 8 Math Previous Results	 New Results	Grade 12 Math Previous Results	 New Results	Grade 12 Geometry Previous Results	 New Results
GENDER BETA COEFFICIENT						
Intercept	(-0.60)	-0.06	(-2.52**)	-3.92**	(6.91**)	-5.18**
Percent black		-0.81		0.40		0.89
Percent Hispanic		-1.76†		-0.20		-0.08
Average SES	(-1.58)	-2.38*		1.63		2.56
Average time spent on:						
Working in small groups		-0.60		.71		0.54
Working with objects		-0.68	(1.71)	2.67*	(2.71†)	2.98†
Doing problems on worksheets		-2.25		1.07		0.87
Doing problems from textbook		-1.37		-0.94		-1.08
Taking math tests		-1.26	(-2.88†)	-2.86†	(-1.63)	-3.64†
Using computer		-0.87		-1.39	(-3.21*)	-2.50
Using calculator		0.56		-0.42		0.36
Writing math proofs				-0.64		-0.96
Formulating own math problems				1.42		2.21
RACE-ETHNICITY BETA COEFFICIENT						
Intercept	(-16.00**)	-18.14**	(-14.88**)	-15.64**	(-13.99**)	-16.37**
Percent black		0.67		-0.49		-2.02
Percent Hispanic		0.09		0.20		-.36
Average SES	(-1.81)	-3.91**		1.81		2.09
Average time spent on:						
Working in small groups		2.17		-0.37		-1.28
Working with objects	(3.16*)	3.62*		2.08		2.35
Doing problems on worksheets		-0.68		2.56		3.26
Doing problems from textbook		-2.32		-5.40		-3.57
Taking math tests		-1.34		-3.73	(-6.71†)	-5.35
Using computer		-0.84		-2.93		-3.42
Using calculator		-1.84		-1.51		-2.93
Writing math proofs				-0.53		-0.72
Formulating own math problems				1.45		3.40

[1]All between-school independent variables have been standardized. See technical notes for more information.

NOTE: **probability ≤ .01; *probability ≤ .05; †probability ≤ .10. This table shows only the gender and race-ethnicity parameter equations of the estimated HLM model. The complete HLM model includes equations predicting the intercept. SES is fixed at the student level and not modeled at the school level.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 8th Grade and 12th Grade Students, Restricted-Use Data Base.

Table 6.4.—Parameter variance statistics for classroom instructional methods predictors of student-level parameters of grade 8 math, grade 12 math, and grade 12 geometry, after coursetaking has been removed from the within-school equation

PARAMETER VARIANCE STATISTICS	Grade 8 Math	Grade 12 Math	Grade 12 Geometry
GENDER BETA COEFFICIENT			
Parameter variance	33.34	14.89	23.06
Chi-square	179.81	157.76	148.27
Degrees of freedom	137	124	124
Probability value	.017	.029	.101
RACE-ETHNICITY BETA COEFFICIENT			
Parameter variance	·9.14	49.56	94.64
Chi-square	185.58	155.35	161.02
Degrees of freedom	137	124	124
Probability value	.006	.032	.024

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 8th Grade and 12th Grade Students, Restricted-Use Data Base.

In both grades, the model predicting the intercept, or average achievement, estimated the same equation as before, while the models predicting the gender and race-ethnicity gaps produced some new results in the predictors (table 6.3 compared to tables 4.9, 4.10, 5.9, 5.10). However, the gender and race-ethnicity intercepts in these equations were very similar to the models in which coursetaking was included. The gender gap remained the same in grade 8 math, while it increased by only 1 point in grade 12 math, and decreased by only 1 point in grade 12 geometry, indicating that coursetaking had little effect. However, the race-ethnicity gap consistently increased by 1–2 points in grade 8 math and grade 12 math and geometry, implying that coursetaking was capturing a small amount of the gap.

In grade 8, there were two new results in the independent variables. The higher the average SES of a grade, the lower the math achievement of girls in relation to boys and the lower the math achievement of African-Americans, Hispanics, and Native Americans in relation to whites and Asian-Americans. Therefore, in higher SES schools, both the gender and race-ethnicity gaps in 8th grade math were larger. These differences may be due to coursetaking since this association was not present when coursetaking was controlled for. In addition, the previous result that working with objects predicted a smaller race-ethnicity gap in grade 8 math remained significant.

In grade 12, there was one new result in the math equations and one new result in the geometry equations. In the model predicting the gender gap in math, in schools where 12th graders worked with objects more than the average amount, the gender gap was lower by about 3 points, which almost equaled the 4 achievement points that females averaged less than males in grade 12. This variable had not been significant in the models that included coursetaking. Since the gender difference was present whether or not coursetaking was controlled for, but the association with objects that counteracted that gender difference appeared only when coursetaking was not controlled for, this method may be related in some way to coursetaking. This relationship needs to be investigated to determine whether it is the working with objects or taking the courses they are associated with that predicts more equal achievement for females in relation to males.

In the model predicting the gender and race-ethnicity gaps in 12th grade geometry, no new variables became significant. However, one variable became almost-significant, one variable that had been almost significant became nonsignificant, and one other that had been significant no longer was. In the gender gap equation, the significant relationship between the use of computers in math classes with the gender gap became nonsignificant, while the nonsignificant relationship between math tests and the gender gap became almost significant. In the race-ethnicity gap equation, the almost-significant negative relationship between taking math tests and the race-ethnicity gap also became nonsignificant. Thus, the presence of the coursetaking variables subtly shifted the relationship of these variables to the gender and race-ethnicity gaps.

Chapter VII

Summary and Discussion

A. Overview

Using HLM on the 1990 NAEP mathematics data in grades 4, 8, and 12 produced several interesting results with models that attempted to explain variations in average math achievement, the gender gap, and the race-ethnicity gap. The results of these models are summarized and discussed in this chapter. After analyzing the meaning of the results and the questions that still need to answered, the discussion then turns to evaluating the lessons learned about the HLM methodology and NAEP in the process of analyzing these NAEP data. Finally, recommendations for further research and for improving the use of both NAEP and HLM are provided.

B. Models predicting average achievement

This summary of the results is guided by two questions. First, how do the results on the 1990 NAEP data compare to similar HLM models tested on the 1986 NAEP data in grades 3, 7, and 11? Second, what are the results and the interpretations of the results of the new models? The comparisons between the old and new results are presented first, and then the interpretations of the new results are discussed.

Models tested on both 1986 and 1990 NAEP data

Only two models—student body characteristics and school resources—were similar to the HLM models tested on the 1986 NAEP data. For the most part, the results of these models differed from the earlier similar HLM models. The only result that closely replicated findings from the earlier study was in the student body characteristics model, where schools with higher percentages of African-American students had lower average achievement in each grade. The 1986 and 1990 models estimated a similar lower level of achievement for grade 3 in 1986 and for grade 4 in 1990. In 1990, the average achievement was 5 points lower for each standard deviation above the average percent African-American, while in 1986 it was 8 achievement points lower. The grade 7/8 and 11/12 results in 1986/1990 are almost identical (11 and 12 achievement points lower in grades 7 and 8, respectively, and 9 and 10 achievement points lower in grades 11 and 12, respectively). These results point to a persistent lower achievement in schools with higher percentages of African-American students compared to other schools, irrespective of student body SES levels (and in 1990, controlling for coursetaking in grades 8 and 12). This association needs to be investigated further to identify the factors involved in the relationship between African-American students, their schools, and achievement that may be operating at every SES level.[65]

In both 1986 and 1990, the direction of the association of school SES with achievement was the same. Schools with higher average SES levels averaged higher

[65]One factor that may be eliminated is test bias in the assessment instruments. NAEP assessment items undergo extensive analytical and judgmental reviews for test bias by race-ethnicity or gender.

achievement. However, the size of the effect varied by grade between the years. In 1986, the effect of SES was large in grade 3, and although significant, was almost negligible in the higher grades. In contrast, in 1990, the effect of SES levels was larger in the higher grades.

Part of the difference in results may be due to the different measures of SES of the student body that were used in 1986 and 1990. In 1986, school-level variables concerning the percentage of students in the school in the Chapter One and free lunch programs made up a school-level "disadvantaged" index, which expressed the economic situation of students in the entire school. In 1990, student-level SES variables of parents' education and reading materials available in the home were aggregated by grade to the school level. These variables expressed a type of educational SES level, averaged for students in each grade for each school. It is possible that the school-wide economic variables used in 1986 were salient only for students in the early grades, and in fact, many schools in the sample did not participate in the lunch program in the later grades. In addition, the 1990 variables that made up the SES value were all conditioning variables. The use of conditioning variables, although aggregated, may have allowed an association between student body SES levels and achievement to appear.

Unlike 1986, the percentage of Hispanic students had little or no association with average achievement in 1990. Whereas in 1986 there was a 5-point lower average achievement for every standard deviation above the average percent Hispanic in each grade, in 1990 there were no significant difference in grades 4 and 12, and only a slight difference in grade 8. Since the same measure was used in both study years, it is possible that whatever effect was there had been reduced in four years. However, the difference in results could also be due to random year to year fluctuations that cannot be detected with just two time points.

In 1990, the only school resource associated with achievement was the percentage of students using computers in grade 4. This computer variable could indicate that the use of computers is related to math achievement in grade 4, or it could mean that schools with more students using computers are also providing other resources or instruction that are associated with higher math achievement.

These 1990 results are quite different from those in 1986, when there were several more significant school resources variables. In 1986, district instructional funds per student were positively associated with achievement in grades 7 and 11, student size was positively associated with achievement in grades 3 and 11, and having a specialized science lab was positively associated with achievement in grade 11. Having a specialized science lab was not estimated in 1990 because it was available only in grade 4. However, it seems puzzling that district instructional funds was not significant in 1990 as it had been in 1986. However, when it was significant in 1986, only 2 average achievement points were gained for every standard deviation above the mean amount of funds, so maybe this small effect occurred by chance only in 1986. It was not a conditioning variable in either year, so it is also possible that something more inherent in the relationship between math achievement and the funds available may have changed. It is even more puzzling that a larger school size would not predict higher achievement. In 1986, the model predicted an average of 2 to 4 points higher for grades 3 and 11 for every standard deviation above the average school size, while in 1990 the differences were under 1.5 and not significant. It is possible that new variables introduced in 1990 accounted for the earlier school size effect. However, while these differences could be important changes, more time points are needed to determine whether they are actual trends or random fluctuations.

New models tested only on 1990 NAEP data

The other four models tested variables that not only were different from those in the earlier study, but were conditioning variables. In addition, they were more directly related to instruction and the school climate for learning than models in the earlier study. All of these factors made them more likely to be significantly related to average achievement, and there were more significant results.

In the classroom instructional methods model (derived from student self-reports), most instructional methods that were significantly related to average achievement confirmed common knowledge about effective instructional methods. However, some results were surprises and need to be investigated further.

Working on problems from textbooks was associated with higher achievement in math and geometry in both grades 4 and 8, while working with objects was associated with higher geometry achievement in grade 4 only. These results could either show that these are effective learning methods, or indicate that schools that use these methods may have other characteristics that support higher achievement. These characteristics would need to be identified using a more qualitative style of research. The association of working with blocks, rulers, and shapes with higher 4th grade geometry achievement supports the theory that younger children learn math and geometry best when they can manipulate objects that illustrate math concepts.

In grade 4, using calculators was associated with lower math achievement. In 8th grade, using computers was negatively associated with math and geometry achievement. However, in 12th grade, using calculators was positively associated with math and geometry achievement. Using calculators and computers during math instruction has come under increasing scrutiny, and results here provide some evidence that they might have different meanings for different age groups. In grade 4, using calculators in class might be detrimental, since they are associated here with lower achievement. Using calculators may be taking time away from basic instruction, or students might be using them instead of solving arithmetic problems on their own, and thus not learning those skills or concepts. Alternatively, the test may penalize students who usually use calculators and may be testing more routine applications than higher level concepts.

Similarly, in grade 8, the use of computers was associated with lower achievement. Either computers are not useful for this age group, or the way they were used by these schools was associated with lower rather than higher achievement. It is possible that computers were used for remedial instruction and that they indicate, rather than cause lower achievement. By grade 12, calculators may be used to support more higher level applications rather than as substitutes for knowledge of basic math, so their use may indicate the presence of higher level learning. However, more information is needed about how calculators and computers are being used in these classrooms before any definite conclusions can be made.

In grade 12, writing proofs was positively associated with math and geometry achievement. Since it is in geometry that most math proofs are written at this level, these results suggest that students in successful geometry classes write more math proofs. However, it is unclear why formulating one's own math problems would be negatively associated with geometry achievement. This result could either reflect a problem with this method, or with the definition of the variable.

The school climate models showed that attitudes that students hold about math are related to math and geometry achievement. In both grades 8 and 12, schools with more students who liked math and felt that they were good at math averaged higher math achievement, and 12th graders in these types of schools averaged higher geometry achievement. It is possible that this result did not show up in grade 4 because there was less variation in these attitudes between schools than in grades 8 and 12. In grade 12, however, schools in which students believed math was useful averaged lower math and geometry achievement. This may reflect the number of students taking vocational or applied math as opposed to higher level courses such as geometry and calculus.

One important result for grades 4 and 8 was present in the school climate model that controlled for academic expectations of a school. Schools that supported a more female-positive atmosphere, that is, had more students disagreeing that math is more for boys than girls, averaged higher levels of math and geometry achievement in grade 4 and math achievement in grade 8. This relationship was also present for 4th grade math when controlling for problems in the school. While this attitude did not narrow the gender gap in any grade, it is significant to note that attitudes that support females actually may support all students. This result is even more remarkable because 85 percent of the schools in grade 4 and 8 (and 81 percent of the grade 12 schools) had female-positive atmospheres. Therefore, the few schools with female-negative atmospheres had to average relatively low achievement to distinguish themselves from the female-positive schools, and conversely, the female-positive schools must have averaged relatively high achievement. It is not possible to tell which was the cause and which the effect, or whether some third variable caused both of these factors, but high achievement and female-positive atmospheres are associated with each other in these schools.

The school climate model with student behavior and safety issues pointed to only one behavioral variable that was associated with achievement. In grade 12, classroom disruption by students was related to achievement. Although higher incidents of classroom disruption and safety threats were measured in grade 8, they do not seem to be associated with achievement until high school. A low average number of problems in the school was recorded in grade 4, so these issues must not be as salient in that grade.

The school climate model with academic expectations used different measures than the 1986 "academic press" model. Of the many variables in the 1986 study that measured academic press, such as academic rigor, changes in school academic standards, and parent/school interactions, only one—the changes in the standards—was significantly associated with achievement, and it predicted lower achievement in grade 7. This "change" variable seemed to be identifying which schools needed to change rather than ones that had improved dramatically. Consequently in 1990, different variables were chosen to represent academic pressure that were more closely related to the academic climate and expectations for math achievement that surrounded the students in each grade, such as time in math instruction and amounts of math taken by students in each school.

Using these variables, which were also conditioning variables, did produce the expected results for grades 8 and 12. The more 8th grade students who were taking algebra, and the more 12th grade students who were in the academic track and had taken calculus, the higher the average math and geometry achievement in a school. However, it is not possible to tell whether these variables are actually measuring academic pressure itself or whether they represent the effect of the pressure on the students. In either case, it is clear that having more students in higher level classes and programs is associated with higher achievement. The grade 4 variables of amount of math instruction and the priority of math in a school were not significant. However, they were also not conditioning variables, so it is hard to tell whether they were not significant because of that reason, or because they truly had no relationship to achievement.

C. Models predicting the gender and race-ethnicity gap

Gender

In the study of 1986 NAEP data, there were only two significant factors associated with the gender gap in math achievement. These two factors were that the higher the proportion of African-American students in an elementary school, the better 3rd grade females did in relation to boys; and the larger the school size in a high school, the better 11th grade females did in relation to boys. It was assumed that more variables were not significant in that study because either variables were not chosen that might specifically relate to gender differences in achievement or none of the school-level variables were conditioning variables.

For the present study, the classroom instructional methods and school climate models contained the variables in NAEP that most closely addressed the subtle issues of instructional methods, classroom interactions, and attitudes toward math that are theorized to create or reinforce gender differences. These variables were specifically expected to relate to gender differences, especially in grades 4 and 8 where there was variation in the gap between females and males. In addition, these variables were all conditioning variables. However, only one variable, in one grade and one subject, was associated with the gender gap. In grade 12 geometry, the amount of time spent using computers in math class predicted a larger gender gap. After discussing this one result, the reasons for the lack of other results will be reviewed.

The gender gap in grade 12 geometry was wider in schools in which computers were used more in 12th grade math classes. This negative association between gender and computers is of concern because experience with computers can be as important as math skills as a basis for success in technological fields. Therefore, the reasons for this association need to be investigated with more school and classroom-based research, guided by the following questions. Some of these questions can be answered using the NAEP dataset, while others can only be answered through observational research in schools. The questions include: How are the computers used? Do the computers have anything to do with learning geometry skills, or are they proxies for some other sort of access to instruction? Is there a gender difference in computer use, or in access to math or geometry teachers or facilities? Are there as many females as males taking geometry?

No school climate variables were associated with the gender gap in grades 4, 8, or 12. Interestingly, the presence of female-positive attitudes was not related to the gender gap, although it had been associated with higher achievement for all students in grade 4 and 8.

In general, most variables were not associated with the gender gap. It could be that the factors associated with gender differences are too subtle to be picked up by the measures available in NAEP. The problems with the variables discussed in chapters one and two may also contribute to a lack of results.

In addition, there may not be a strong relationship between present classroom and school factors and present achievement by gender in the earlier grades because the latter are largely the result of past factors. Instead, the effect of early classroom and school factors may be more of a lagged effect that causes average gender differences to appear in high school rather than earlier. Alternatively, factors that may cause gender differences in math may not start until adolescence. This theory would fit with studies that show that it is after puberty that females start to underachieve in math in relation to males due to social pressure and fear of competing with males. In any case, the process of channeling females into lower level math achievement may be both subtle and complex, and this process may be hard to measure in a cross-sectional study. In addition, the lagged development of females and males in math may proceed at different rates, and differences between females and males may appear at different times. In this case, a cross-sectional study could pick up the differences but not be able to explain them.

Another explanation for so few gender differences and so few predictors of the gender gap is that in grades 8 and 12, coursetaking was controlled for at the student level. Since gender differences in math achievement are often first manifested as course-taking differences, and it is those course-taking differences that lead to actual achievement differences, controlling for this variable may have removed the effects of, as well as the explanations for, channeling females into lower math achievement. However, the exploratory analysis in chapter six on models with coursetaking removed showed that this change yielded only a slight change in the race-ethnicity gap and little change in the gender gap. Therefore, coursetaking is not an explanation for the lack of results, and the gender differences and the one predictor of the gender gap discussed here are valid regardless of amount of coursetaking.

Race-ethnicity

Differences in math and geometry achievement by race-ethnicity were much more pervasive than differences by gender. The gap between African-American, Hispanic, and Native American students and white and Asian-American students consistently averaged 14 points in grades 4, 8, and 12, even controlling for SES level. In addition, this gap varied between schools in grade 8 and 12, and almost varied in grade 4. This suggests that factors in schools have the potential to explain math achievement differences associated with race-ethnicity.

Two variables in the school resources model were related to the race-ethnicity gap in grade 12. Schools with higher student/teacher ratios and more district funds per student had substantially lower gaps in 12th grade math achievement between 12th grade African-American, Hispanic, and Native American students and white and Asian-American students. These factors are somewhat contradictory, but they can both be manipulated by school districts. It is not clear why having more students per teacher would help the achievement of African-American, Hispanic, and Native American students in relation to white and Asian-American students, so this relationship needs to be investigated further before drawing any policy conclusions. Perhaps funds available due to larger class sizes were used for classes with high percentages of African-American, Hispanic, and Native American students. Increasing district funds per student may seem to have a more obvious connection to increasing African-American, Hispanic, and Native American achievement, but even if the relationship is not spurious, it would be important to know what the funds were spent on and whether they actually had a direct effect on math instruction.

Variables in the classroom instructional methods model also pointed to factors associated with the race-ethnicity gap that may be amenable to educational intervention, if confirmed by further research. The race-ethnicity gap in grade 4 math and geometry was smaller in schools where students worked on worksheets more. This could show that this instructional method is better suited for equal learning in this age group, or it could be a proxy for a type of classroom atmosphere that is maintained when students spend more time on worksheets. This atmosphere, such as more order in the classroom, more controlled instruction from the teacher, or more monitored interaction with other students, may counteract differences by race-ethnicity. Only ethnographic or experimental research could confirm the reasons for this smaller race-ethnicity gap with these methods.

In grade 8 math, the race-ethnicity gap was smaller in schools where 8th grade students worked more with objects. Learning math by manipulating objects has long been known to help students in the early grades, and this result may show that this method plays a particular role in helping African-American, Hispanic, and Native American students in grade 8 to catch up in achievement to the achievement levels of white and Asian-American students. However, it does not explain why white and Asian-American students would not also benefit so much from this method. Again, this method may represent a type of classroom atmosphere that fosters productive learning for those students who most need it, and somehow counteracts any differences by race-ethnicity that might already exist.

D. Lessons learned about using HLM with NAEP data

The effectiveness of the HLM models with NAEP data

The outcomes of this study were the result of interactions between the type of data available in NAEP, the types of models that can be tested using HLM, and the way these interactions were expressed in the HLM estimates and statistics. Both NAEP and HLM worked very well when explaining variations in achievement. However, they did not work particularly well in explaining variations in the gender and race-ethnicity gaps. The reasons for this difference seem to stem from limitations of both NAEP and HLM, and these reasons are discussed below. In addition, in the process of analyzing these results, two patterns appeared that are also related to characteristics of both NAEP and HLM. The first pattern was a lack of correlation between the amount of parameter variance, reliability, significant variables, and the proportion of variance explained. The second pattern was a sensitivity of the models to slight changes in variables. These are also discussed below.

Explaining the wealth of results for average achievement

The HLM models were very successful in explaining the variations in the average NAEP achievement data, at least in terms of proportion of variance explained. This is not surprising given that the assessment scores, even in the form of plausible values, are the product of years of refinement by NAEP, and are the dataset's "best" variables. In addition, the intercept equation in an HLM model is usually the most reliable, because the dependent variable usually has the least sampling error and the most parameter variance. In fact, HLM statistics showed that this variable had high reliability and high parameter variance.

While the three student body characteristic variables were able to explain a majority of the variance in grades 8 and 12, several other variables were also significantly related to achievement. Some of these findings were robust across models, and others wavered on the edge of being significant. Any of these results may be affected by the problems with variable specification mentioned in chapters one and two. However, these models were still successful by any HLM measure, and more testing can be done to determine whether these particular variables made a difference to that success, or whether any three variables would account for that much variance.

Explaining the lack of results for the gender and race-ethnicity gap

The HLM models were less successful in explaining the gender and race-ethnicity gaps. In the case of the gender gap, this may have been simply due to the small amount of gender gap in the early grades, and the small amount of variation in the large gap in grade twelve. However, there was enough gender gap and variation in all grades at some point to be discouraged about the lack of significant predictors. In addition, the exploratory analyses in chapter six found that neither taking out course-taking nor fixing the race-ethnicity coefficient changed the gender gap results substantially. The race-ethnicity gap, which was large, significant, and variable, also had few significant predictors. Therefore, other reasons for these results will be considered.

First, the other coefficients besides the intercept in HLM models are usually harder to predict. This is due in part to being a more indirect measure of the dependent variable, since they incorporate information about both achievement and gender or race-ethnicity. This gives them more sampling error and less parameter variance. In fact, the reliabilities of the gender and race-ethnicity coefficients were very low, which made even a high amount of parameter variance explained account for little of the total variance. Second, with less reliable dependent variables, the models might be more sensitive to any problems with the predictor variables. Although the student body characteristics, classroom instruction, and school climate variables were theoretically very related to gender and race-ethnicity, their actual measurement with the NAEP data may have not been precise enough to capture the true associations. In addition, these models would also be more sensitive to the problems in using school-level classroom data and other aggregated variables because the gender and race-ethnicity gaps are, like these variables, school-wide aggregates of individual and classroom patterns.

NAEP and HLM can continue to be used to explore the gender and race-ethnicity gaps. Besides constructing and testing other predictor variables, variations of the gender and race-ethnicity gaps can be created by testing the interaction terms or different combinations of gender and race-ethnicity. However, if even these simple models were not successful, it is possible that the limitations of the data and the model may prevent the combination of HLM models and NAEP data from obtaining more meaningful results in this area.

Inconsistency in the relationship between reliability, parameter variance, and significant variables when examining separate school-level equations

In the average achievement models, the HLM statistics were fairly consistent. The reliability was high, the parameter variance was high, some variables were significant, and a high proportion of parameter variance was explained. There was a tendency for the reliability to drop as the proportion of variance explained rose in grades 8 and 12, which was probably due to new variables that added sampling error. With different variables, grade 4 reliability remained high and a low proportion of variance was explained. For the most part in grade 4, the reliability and parameter variance were high, and some significant variables did account for some of the variance. Still, like in OLS regression, the presence of significant variables does not determine how much of the variance is explained, and the school variables could not explain as much of the grade 4 variance as the variables could explain of the grade 8 and 12 variance. Some of the variables differed by grade, so it is possible that the particular variables simply had less salience in grade 4 than the higher grades.

However, in the gender and race-ethnicity models, the HLM statistics were much less consistent. The following situations illustrate the inconsistency of outcomes in these statistics. Despite the presence of parameter variance in the grade 4 and 8 gender gap in math, no significant variables were found and no variance was explained by the student characteristics model, and significant variance remained. This pattern was expected. However, after controlling for this model, significant parameter variance in grade 12 appeared after previous nonsignificance. In a more dramatic example, grade 12 math and geometry had similar low reliabilities in the instructional methods model (.12 and .14). In grade 12 geometry, there were no significant variables, yet 37 percent of the variance was explained, despite almost no variance explained by earlier models with only student characteristics. However, in grade 12 math, there was one significant variable, and 37 percent of the variance was also explained. Both math and geometry in this case had one or two variables that were almost significant, but neither would account for that much variance explained. In some cases, significant predictors seemed to account for more of the variance, but in other similar models, some of the variance was explained even with no significant predictors. There were also situations in which the proportion of variance explained dropped in later models, despite having the same student body variables in the model.

These seemingly inconsistent results can be explained by factors in both HLM and NAEP. First, although the school-level equations for average achievement and the gender and race-ethnicity gaps were presented separately in this report, they are all part of an interconnected HLM equation for each model, as shown in the Appendix A tables. The "separate" equations are not separate at all, but are greatly affected by the other Betas, their parameter variances, and their school-level predictors in the other equations. Therefore, what seems inconsistent in one equation is actually a meaningful part of a larger HLM equation. In short, although these equations can be examined separately, their interconnectedness to the rest of the HLM model must not be forgotten, and none of these HLM statistics can be given the same interpretation as similar OLS statistics.

An additional explanation for some discrepancies in the parameter variance patterns is the use of plausible values in NAEP. Each parameter variance reported in this study is the average of five parameter variance values. When these values are compared to determine proportion of variance explained, fluctuations in the averages sometimes cause the unconditional parameter variance to be smaller than the variance in later models, or for more variance to be explained by earlier than later models. For this reason, small changes in the proportion of variance explained are not as important to consider as larger trends.

Sensitivity of the models to specific variables

The six models tested seemed somewhat sensitive to the particular variables in each model. This could be observed in the school climate models, which contained the same school climate control variables but had control variables that were significant in only one of these models. For example, the percentage of Hispanic students in a school was only associated with the race-ethnicity gap in the last school climate model. Although it had been almost significant in the first school climate model, it was never significant in earlier models. Therefore, it seemed sensitive to the variables in the model.

Identifying the variables that were almost significant provided a larger sample of variables to track, and it was found that some variables, like percentage of Hispanic students, wavered on the edge of significance. While the purpose of multivariate analysis is to identify the correctly specified model that will account for most of the variance and express the true significant relationships, it is hard to identify important associations when a variable behaves differently in different models. While this may be another result of the complex HLM model, it makes variable specification and choice even more important.

Interpreting HLM and NAEP results

Despite the problems of variable specification, lack of significant predictors, inconsistency in the statistics, and sensitive models, it is still possible to produce meaningful interpretations of HLM/NAEP results. First, the more the construction and univariate characteristics of a variable are known, the easier it is to explain their association with other variables. Second, variables that are significant across varying models with different control variables are more robust predictors than those that only appear in one model. This is also a good reason to combine models for a final test of the major variables. Third, while statistical significance is necessary for interpreting a result, it may not always be sufficient. The practical significance of the results must be considered also. Given that the four anchor levels in math are 50 points apart, and the standard deviations around average achievement are between 30 and 40 points, variables that predict average achievement or gap differences of under 5 points may not be as important as those that predict changes of 10 or more points.

The use of conditioning variables in HLM models

Although the 1990 study was designed to determine whether using more conditioning variables would make a difference in the results and whether the conditioning variables would perform better than the nonconditioning variables, these questions were only partially answered.

More results were obtained using the 1990 dataset than in the previous study using the 1986 dataset, although different variables were used in four out of six models. Table 7.1 displays all the student- and school-level variables, and shows which ones are conditioning variables and which ones had significant relationships with student achievement. Most, but not all, of the significant variables were conditioning variables. Many of the conditioning variables that were significant were student-level variables that had been aggregated to the school level. Since they had been used as conditioning variables at the student level, it is interesting that they were associated with within-school parameters after aggregation. The use of conditioning variables is most important at the student level, which is the level at which the plausible values are estimated. However, they still might have some effect at the student group mean (i.e., school level), and these results might support that conclusion.

Using conditioning variables in these models seems to have contributed to finding more significant associations with average achievement, although it did not help explain the gender or race-ethnicity gap. Although most of the conditioning variables were student-level variables aggregated to the school level, many were significant and had moderate effect sizes. This suggests that one possible reason for the lack of results in the 1986 study was that few conditioning variables were used, and also suggests that conditioning variables should be used whenever possible, even if they are aggregated to the school level from the student level. However, since most of the conditioning variables used in 1990 were not available or not tested in models in 1986, it is not possible to tell whether their significance in 1990 is actually due to the fact that they are conditioning variables. In addition, it is important to remember that while conditioning variables may be more likely to be significant, they are not always significant. However, their presence may make it more likely to find significant results.

Table 7.1.—Significant results of conditioning and nonconditioning student-level and school-level variables, by grade: 1990

Variable	Significant Variables		
	Grade 4	Grade 8	Grade 12
Student-level variables			
Gender (1=female)			**
Race-ethnicity (1=Afr. Am, Hisp., Nat. Am.)[1]	**	**	**
SES level	**	**	**
Taking algebra in grade 8 (0/1)		**	
Number of years of geometry by grade 12 (0-3)			**
Number of years of calculus by grade 12 (0-3)			**
Number of students	5,080	5,198	4,953
School-level variables			
Student body characteristics			
Percent black (0-100)	**	**	**
Percent Hispanic (0-100)			
SES level	**	**	**
School resources			
Number of students			
Student/teacher ratio			
Instructional funds/student (1-9)			
Microcomputers/student			
Percent students using computer (0-100)	*		
Classroom instructional methods			
In math class, how often:			
Work in small groups (1-5)			
Work with objects (1-5)	*		
Do problems on worksheets (1-5)			
Do problems from textbook (1-5)	*	*	*
Take math tests (1-5)			
Use calculator (1-5)	*		*
Use computer (1-5)		*	
Write math proofs (1-5)			**
Formulate own problems (1-5)			
School climate: Math attitudes			
Math is useful (-1-+1)			*
Enjoy and feel competent in math (-1-+1)		*	*
Math not more for boys (0-1)	*	*	

Table 7.1.—Significant results of conditioning and nonconditioning student-level and school-level variables, by grade: 1990—Continued

Variable	Significant Variables		
	Grade 4	Grade 8	Grade 12
School climate: Student behavior and safety			
Problems in the school (0-3)			
Percent enrolled full year (0-100)			
Average days absent in month			
Students feel classes often disrupted (1-4)			*
Students feel unsafe at school (1-4)			
School climate: Academic expectations			
Instruction/week in math (1-5)			
Math special priority (0/1)			
Mean composite math score			
Percent taking algebra in grade 8 (0-1)		**	
Percent of grade 12 in academic/college track			**
Average number years of calculus by grade 12			**
Number of schools	257	174	186

[1]African-American, Hispanic, and Native American students were compared to white and Asian-American.

NOTES: Conditioning variables and their results are shown in **bold**. ** probability ≤ .01; * probability ≤ .05.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th, 8th, and 12th Grade Students, Restricted-Use Data Base.

E. Recommendations for further research, use of HLM, and NAEP

Further research

This study is only the beginning of a detailed analysis of the relationships between gender, race-ethnicity, and achievement using HLM and NAEP data. For example, only main effects were tested in this study in order to keep the interpretations as clear as possible. However, the next stage of this research could examine the interaction of race-ethnicity with both SES and gender using either interaction terms or separate models for subgroups. For instance, instead of predicting average achievement in schools, examining the average achievement of the female and male students separately might be a more productive way to measure and predict gender differences.[66] In addition, further research can investigate such areas as the associations between SES and attitudes towards math and between attitudes towards math and course-taking patterns.

[66]Sample sizes are too small to test a separate HLM model for most race-ethnicity subgroups.

This study also identifies areas that can be further investigated using more qualitative research methods. For instance, the associations found between types of instructional methods and the gender and race-ethnicity gaps need to be confirmed and explained in observational studies. Furthermore, the complex relationships between SES, gender, race-ethnicity, instruction, and attitudes towards math that result in the channeling of female, African-American, and Hispanic students into lower level math courses and achievement need to be identified and described using a combination of quantitative and qualitative methods.

Limitations on research questions using HLM

HLM allows researchers to ask the question, what school-level characteristics are associated with gender and race-ethnicity gaps in math achievement? If more of these school factors can be identified with HLM and then qualitatively shown to have a direct affect on achievement, then these factors can be altered within schools. To this end, HLM allows analysts to use correlational tests to identify the types of schools that have a higher or lower gender or race-ethnicity gap in achievement. However, only qualitative or experimental research can identify any causal relationships.

In the future, large datasets may be designed as evaluation and/or experimental studies in to answer the question: What is the effect of a particular intervention technique on either average achievement or on the gender or race-ethnicity gap? Unfortunately, HLM cannot be used to evaluate individual programs or schools because the model "borrows" information from all schools and then "shrinks" the estimates of the smaller schools towards the means of larger, similar schools.[67] Therefore, appropriate caveats should be included in all HLM research that might have causal or evaluative implications.

Recommended changes to NAEP

In many ways this study came up against the limits of the use of NAEP data for studying school-level correlates of achievement and of gender and race-ethnicity differences in achievement. While many excellent indicators were included in NAEP, their presence led to a desire for more and even better measures.

First, using a cross-sectional study of students in schools to find the association of school or classroom factors with student achievement limited the meaningful association, if not the statistical association of these variables. In order to relate classroom activities, teacher characteristics, and student achievement while controlling for previous achievement, within-classroom samples are needed. In addition, academic growth during a school year needs to be measured. Being able to predict improvement from the beginning to the end of a school year or from Spring to Spring would provide more accurate dependent variables. Although testing twice a year would be prohibitive for NAEP, the recording of students' previous test scores or grades might be possible in order to have at least one measure of achievement gains relative to other students from year to year.

If classroom samples are possible, additional variables that are more closely tied to instruction are needed. Reducing the level of missing values in the classroom-level variables and adding other variables reported by teachers would be a good start to measuring the ranges in types of instruction. In particular, while classroom instructional

[67]See the review of this and other limitations of HLM in C.L. Arnold, "An Introduction to Hierarchical Linear Models," *Measurement and Evaluation in Counseling and Development* 25 (2)(July 1992): 83-84.

methods are hypothesized as key variables in how females and minority students learn, they need to be measured at a more specific level, including such factors as the amount of time that a student discusses math problems. In addition, more information is needed about the teachers' interactions with each student, and the student's interactions with other students in groups and during other instructional methods.

Although classroom samples and interactions may be beyond the scope of NAEP, other more descriptive information such as the gender and race-ethnicity distribution in classrooms, especially in grade 12 geometry and calculus and in instructional/work groups, would add valuable information. While there were some teacher gender and race-ethnicity characteristics that could have been used for role model availability, in most grades these teacher and classroom-level variables had too many missing values to be useful. Insuring that these variables are not missing, and adding some teacher variables about mentoring and interactions with specific students would make this information more useful.

In order to address the issue of missing values in the teacher variables, either information on the missing values or more successful teacher data collection is necessary. As mentioned in the introduction to this report, it was not clear whether too few teachers had been sampled, too few students were taking math, or too few teachers had responded. Likewise, more information on the selection, derivation, validity, and reliability of the student background, classroom, and school climate questions would be helpful and would contribute to an understanding of what these variables actually measure.

Finally, the use of plausible values, particularly those that have been conditioned with student and school variables, is still somewhat mystifying to researchers using NAEP data to identify student and school correlates of achievement. Although the conditioning procedure has been justified in several technical articles, and the use of conditioning variables was an important aspect to this study, it is a very confusing concept to explain to new users of NAEP. It would be helpful to have a non-technical explanation of how it is possible to create an estimated outcome variable with predictors, put those predictors into regression equations to explain variance on that outcome variable, and expect to find significant results. In addition, there is nothing written on the procedure of aggregating student-level conditioning variables to the school level and the effect of using these as school-level predictors on the results. That procedure was followed in this study with some assurances from statisticians that this was correct, but there was nothing written to prove it.

Appendix A

Supporting Tables

Table A1.—Average within-school parameters of grade 4 math achievement

WITHIN-SCHOOL PARAMETERS	Gamma coefficient[1]	Standard error[2]	t Value[3]	
INTERCEPT (AVG. ACHIEVEMENT)	211.35	1.06	199.96**	
GENDER BETA COEFFICIENT	-1.46	0.79	-1.84†	
RACE-ETHNICITY BETA COEFFICIENT	-14.74	1.00	-14.78**	
SES BETA COEFFICIENT	6.20	0.63	9.78**	
Random within-school parameters	**Reliability[4]**	**Parameter variance (Tau)[5]**	**Degrees of freedom**	**Chi-square test of Tau > 0[6]**
INTERCEPT (AVG. ACHIEVEMENT)	0.90	235.98	217	1905.08**
GENDER BETA COEFFICIENT	0.15	23.03	217	252.98*
RACE-ETHNICITY BETA COEFFICIENT	0.07	21.41	217	254.02†

[1]Average of five Gamma values. See technical notes for more information.

[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Parameter variance divided by total variance. Average of five reliability values.

[5]Average of five parameter variance values.

[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05, † probability ≤ .10.

SOURCE: U. S Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 4th Grade Students, Restricted-Use Data Base.

Table A2.—Student body characteristics predictors of student-level parameters of grade 4 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	212.87	0.95	224.89**
Percent African-American	-5.00	0.87	-5.74**
Percent Hispanic	-0.97	0.86	-1.12
Student body race-ethnicity unknown	7.18	5.92	1.21
Average SES	5.01	0.97	5.18**
GENDER BETA COEFFICIENT			
Intercept	-1.39	0.80	-1.74†
Percent African-American	-0.24	0.80	-0.30
Percent Hispanic	0.14	0.82	0.17
Student body race-ethnicity unknown	-4.37	5.44	-0.80
Average SES	1.39	1.03	1.35
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.76	1.04	-14.14**
Percent African-American	-0.10	1.55	-0.06
Percent Hispanic	0.53	1.13	0.47
Student body race-ethnicity unknown	-5.43	6.02	-0.90
Average SES	0.79	1.24	0.64
SES BETA COEFFICIENT			
Intercept	6.18	0.63	9.77**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.86	166.91	213	1498.54**
GENDER BETA COEFFICIENT	0.15	22.37	213	249.95*
RACE-ETHNICITY BETA COEFFICIENT	0.06	15.97	213	253.03†

[1]Average of five Gamma values. See technical notes for more information.

[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Parameter variance divided by total variance. Average of five reliability values.

[5]Average of five parameter variance values.

[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade Students, Restricted-Use Data Base.

Table A3a.—School resources predictors of student-level parameters of grade 4 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	213.12	0.97	218.79**
Percent African-American	-4.69	0.91	-5.14**
Percent Hispanic	-0.18	0.92	-0.19
Student body race-ethnicity unknown	8.72	6.94	1.26
Average SES	5.07	0.97	5.21**
School size (number of students)	-1.34	0.95	-1.41
Student/teacher ratio	-1.36	0.97	-1.41
Student/teacher ratio unknown	-2.54	9.10	-0.28
District instructional funds/student	-0.68	0.97	-0.71
District funds/student unknown	0.33	8.70	0.04
GENDER BETA COEFFICIENT			
Intercept	-1.56	0.82	-1.90†
Percent African-American	-0.29	0.87	-0.34
Percent Hispanic	-0.21	0.88	-0.24
Student body race-ethnicity unknown	-4.93	6.14	-0.80
Average SES	1.49	1.04	1.44
School size (number of students)	1.14	0.92	1.23
Student/teacher ratio	0.41	0.90	0.46
Student/teacher ratio unknown	1.57	7.73	0.20
District instructional funds/student	-0.47	0.94	-0.50
District funds/student unknown	-0.38	7.10	-0.05
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.69	1.11	-13.26**
Percent African-American	0.45	1.68	0.27
Percent Hispanic	0.63	1.21	0.52
Student body race-ethnicity unknown	-4.60	7.18	-0.64
Average SES	0.99	1.23	0.81
School size (number of students)	-0.61	1.29	-0.47
Student/teacher ratio	0.96	1.06	0.90
Student/teacher ratio unknown	-3.23	9.67	-0.33
District instructional funds/student	-1.34	1.26	-1.07
District funds/student unknown	2.31	8.84	0.26
SES BETA COEFFICIENT			
Intercept	6.20	0.63	9.77**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.86	166.48	208	1493.37**
GENDER BETA COEFFICIENT	0.15	22.24	208	245.72*
RACE-ETHNICITY BETA COEFFICIENT	0.08	21.82	208	252.08*

[1] Average of five Gamma values. See technical notes for more information.
[2] Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3] Gamma divided by standard error. Probabilities based on a two-tailed test.
[4] Parameter variance divided by total variance. Average of five reliability values.
[5] Average of five parameter variance values.
[6] Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 4th Grade Students, Restricted-Use Data Base.

Table A3b.—Detailed school resources predictors of student-level parameters of grade 4 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	213.30	1.04	204.34**
Percent African-American	-4.09	0.95	-4.28**
Percent Hispanic	-0.35	0.92	-0.38
Student body race-ethnicity unknown	8.21	6.93	1.18
Average SES	5.07	0.98	5.16**
School size (number of students)	-1.33	1.03	-1.29
Student/teacher ratio	-1.58	0.97	-1.62
Student/teacher ratio unknown	-2.72	9.02	-0.30
District instructional funds/student	-0.71	0.97	-0.74
District funds/student unknown	-0.23	8.65	-0.03
Computers per student	-0.83	0.90	-0.92
Computers per student unknown	4.02	3.42	1.17
Percent use computers as part of math instr.	2.36	1.04	2.26*
Percent using computers unknown	-9.17	4.25	-2.16*
GENDER BETA COEFFICIENT			
Intercept	-1.45	0.91	-1.59
Percent African-American	-0.62	0.91	-0.68
Percent Hispanic	-0.33	0.91	-0.36
Student body race-ethnicity unknown	-5.12	6.18	-0.83
Average SES	1.65	1.05	1.56
School size (number of students)	1.77	1.01	1.76†
Student/teacher ratio	0.71	0.92	0.77
Student/teacher ratio unknown	1.21	7.70	0.16
District instructional funds/student	-0.27	.095	-0.28
District funds/student unknown	-0.33	7.07	-0.05
Computers per student	1.58	0.82	1.93†
Computers per student unknown	-1.40	3.24	-0.43
Percent use computers as part of math instr.	-1.43	0.92	-1.56
Percent using computers unknown	0.94	3.84	0.25
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.75	1.26	-11.74**
Percent African-American	0.18	1.63	0.11
Percent Hispanic	0.66	1.21	0.54
Student body race-ethnicity unknown	-5.48	7.30	-0.75
Average SES	1.25	1.24	1.01
School size (number of students)	-0.02	1.40	-0.02
Student/teacher ratio	1.28	1.08	1.18
Student/teacher ratio unknown	-3.58	9.51	-0.38
District instructional funds/student	-1.07	1.26	-0.84
District funds/student unknown	2.75	8.80	0.31
Computers per student	1.53	1.02	1.50
Computers per student unknown	-0.30	4.25	-0.07
Percent use computers as part of math instr.	-1.52	1.25	-1.22
Percent using computers unknown	0.89	4.72	0.19
SES BETA COEFFICIENT			
Intercept	6.19	0.63	9.77**

Table A3b.—Detailed school resources predictors of student-level parameters of grade 4 math achievement—Continued

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.86	162.09	204	1472.97**
GENDER BETA COEFFICIENT	0.14	21.82	204	239.29*
RACE-ETHNICITY BETA COEFFICIENT	0.06	16.06	204	244.11†

[1]Average of five Gamma values. See technical notes for more information.

[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Parameter variance divided by total variance. Average of five reliability values.

[5]Average of five parameter variance values.

[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress. 1990 NAEP: Math Assessment of 4th Grade Students, Restricted-Use Data Base.

Table A4.—Classroom instructional methods predictors of student-level parameters of grade 4 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	213.04	0.91	233.07**
Percent African-American	-4.70	0.93	-5.04**
Percent Hispanic	-0.48	0.87	-0.55
Student body race-ethnicity unknown	5.97	5.80	1.03
Average SES	4.83	0.96	5.04**
Work in small groups	-1.31	0.95	-1.38
Work with objects	1.96	0.97	2.02*
Do problems on worksheets	0.05	1.01	0.05
Do problems from textbook	2.74	0.96	2.86**
Take math tests	-1.30	1.02	-1.28
Use computer	0.96	0.88	1.09
Use calculator	-2.12	0.95	-2.23*
GENDER BETA COEFFICIENT			
Intercept	-1.41	0.81	-1.75†
Percent African-American	-0.67	0.90	-0.75
Percent Hispanic	-0.33	0.86	-0.38
Student body race-ethnicity unknown	-4.66	5.53	-0.84
Average SES	1.17	1.07	1.09
Work in small groups	1.23	0.87	1.42
Work with objects	-0.31	0.96	-0.32
Do problems on worksheets	-1.73	0.93	-1.86†
Do problems from textbook	-0.03	0.93	-0.04
Take math tests	-0.26	0.94	-0.28
Use computer	-0.37	0.85	-0.44
Use calculator	0.59	0.96	0.61
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.69	1.04	-14.16**
Percent African-American	0.56	1.55	0.36
Percent Hispanic	1.26	1.18	1.07
Student body race-ethnicity unknown	-3.82	6.40	-0.60
Average SES	0.84	1.20	0.70
Work in small groups	0.47	1.11	0.42
Work with objects	-0.86	1.23	-0.70
Do problems on worksheets	2.63	1.28	2.05*
Do problems from textbook	1.00	1.14	0.88
Take math tests	-0.52	1.27	-0.41
Use computer	-0.97	1.11	-0.88
Use calculator	-0.54	1.39	-0.39
SES BETA COEFFICIENT			
Intercept	6.19	0.64	9.69**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.85	152.19	206	1363.43**
GENDER BETA COEFFICIENT	0.15	23.60	206	243.28*
RACE-ETHNICITY BETA COEFFICIENT	0.06	14.37	206	246.94†

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade Students, Restricted-Use Data Base.

Table A5.—School climate (math attitudes) predictors of student-level parameters of grade 4 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	213.02	0.94	226.89**
Percent African-American	-4.59	0.89	-5.17**
Percent Hispanic	-0.90	0.85	-1.06
Student body race-ethnicity unknown	5.92	5.96	0.99
Average SES	3.71	1.07	3.47**
Students feel math is useful	1.44	1.00	1.43
Students enjoy and feel competent in math	1.26	0.96	1.31
Students disagree that math is more for boys	1.92	1.03	1.86†
GENDER BETA COEFFICIENT			
Intercept	-1.40	0.81	-1.73†
Percent African-American	-0.13	0.81	-0.16
Percent Hispanic	0.05	0.83	0.06
Student body race-ethnicity unknown	-3.07	5.55	-0.55
Average SES	1.11	1.23	0.90
Students feel math is useful	1.31	0.95	1.38
Students enjoy and feel competent in math	0.56	1.01	0.56
Students disagree that math is more for boys	-0.67	0.96	-0.69
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.76	1.07	-13.79**
Percent African-American	-0.25	1.54	-0.16
Percent Hispanic	0.62	1.13	0.55
Student body race-ethnicity unknown	-6.40	6.24	-1.03
Average SES	1.39	1.38	1.01
Students feel math is useful	-1.70	1.15	-1.48
Students enjoy and feel competent in math	-0.87	1.13	-0.77
Students disagree that math is more for boys	0.10	1.16	0.09
SES BETA COEFFICIENT			
Intercept	6.20	0.63	9.80**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.86	163.15	210	1432.77**
GENDER BETA COEFFICIENT	0.15	22.88	210	245.90*
RACE-ETHNICITY BETA COEFFICIENT	0.08	19.95	210	248.93†

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade Students, Restricted-Use Data Base.

Table A6.—School climate (math attitudes and student behavior and safety) predictors of student-level parameters of grade 4 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	213.61	0.97	220.75**
Percent African-American	-4.27	0.90	-4.73**
Percent Hispanic	-0.79	0.86	-0.92
Student body race-ethnicity unknown	4.99	5.96	0.84
Average SES	3.59	1.09	3.29**
Students feel math is useful	1.47	1.01	1.46
Students enjoy and feel competent in math	1.47	0.97	1.51
Students disagree that math is more for boys	2.04	1.03	1.99*
Index of problems in the school	-0.59	0.91	-0.65
Index of problems unknown	-8.30	10.55	-0.78
Percent students enrolled all year	0.64	0.97	0.66
Percent enrolled unknown	0.06	10.24	0.01
GENDER BETA COEFFICIENT			
Intercept	-1.51	0.86	-1.76†
Percent African-American	-0.19	0.83	-0.22
Percent Hispanic	0.14	0.85	0.17
Student body race-ethnicity unknown	-2.99	5.55	-0.54
Average SES	1.24	1.27	0.98
Students feel math is useful	1.14	0.96	1.19
Students enjoy and feel competent in math	0.72	1.01	0.72
Students disagree that math is more for boys	-0.73	0.97	-0.75
Index of problems in the school	0.67	0.88	0.76
Index of problems unknown	5.78	10.88	0.53
Percent students enrolled all year	1.45	0.91	1.60
Percent enrolled unknown	-4.36	10.29	-0.42
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.93	1.15	-12.93**
Percent African-American	-0.20	1.51	-0.14
Percent Hispanic	0.68	1.12	0.61
Student body race-ethnicity unknown	-6.52	6.35	-1.03
Average SES	1.17	1.35	0.87
Students feel math is useful	-1.81	1.14	-1.58
Students enjoy and feel competent in math	-0.89	1.12	-0.79
Students disagree that math is more for boys	-0.03	1.17	-0.02
Index of problems in the school	-0.83	1.18	-0.70
Index of problems unknown	1.30	21.74	0.06
Percent students enrolled all year	0.19	1.38	0.14
Percent enrolled unknown	2.03	21.39	0.09
SES BETA COEFFICIENT			
Intercept	6.19	0.63	9.78**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.86	161.98	206	1428.49**
GENDER BETA COEFFICIENT	0.15	22.44	206	241.69*
RACE-ETHNICITY BETA COEFFICIENT	0.07	18.32	206	246.81†

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 4th Grade Students. Restricted-Use Data Base.

Table A7.—School climate (math attitudes and academic expectations) predictors of student-level parameters of grade 4 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	213.75	0.98	217.80**
Percent African-American	-4.15	0.90	-4.59**
Percent Hispanic	-0.83	0.85	-0.98
Student body race-ethnicity unknown	5.34	5.96	0.90
Average SES	3.75	1.07	3.51**
Students feel math is useful	1.55	1.01	1.54
Students enjoy and feel competent in math	1.43	0.98	1.47
Students disagree that math is more for boys	2.07	1.04	2.00*
Amount of instruction in math	-0.24	0.94	-0.25
Amount of instruction unknown	-4.43	5.53	-0.80
Math identified as a special priority	-0.53	0.97	-0.54
Math special priority unknown	-4.17	4.87	-0.86
GENDER BETA COEFFICIENT			
Intercept	-1.42	0.91	-1.57
Percent African-American	-0.50	0.94	-0.53
Percent Hispanic	-0.02	0.83	-0.03
Student body race-ethnicity unknown	-3.73	5.62	-0.66
Average SES	1.33	1.33	1.00
Students feel math is useful	1.32	0.97	1.36
Students enjoy and feel competent in math	0.69	1 00	0.69
Students disagree that math is more for boys	-0.64	0.99	-0.65
Amount of instruction in math	0.66	0.92	0.72
Amount of instruction unknown	0.21	5.06	0.04
Math identified as a special priority	-1.02	0.86	-1.18
Math special priority unknown	0.57	5.31	0.11
Mean composite math score of grade sample	-0.83	1.23	-0.68
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-15.00	1.19	-12.60**
Percent African-American	-0.60	1.64	-0.37
Percent Hispanic	0.53	1.12	0.47
Student body race-ethnicity unknown	-7.66	6.25	-1.22
Average SES	1.28	1.50	0.85
Students feel math is useful	-1.98	1.18	-1.67†
Students enjoy and feel competent in math	-0.78	1.16	-0.67
Students disagree that math is more for boys	0.13	1.20	0.11
Amount of instruction in math	1.88	1.11	1.69†
Amount of instruction unknown	1.78	6.53	0.27
Math identified as a special priority	-0.32	1.19	-0.27
Math special priority unknown	2.53	6.23	0.41
Mean composite math score of grade sample	-0.36	1.68	-0.22
SES BETA COEFFICIENT			
Intercept	6.17	0.63	9.77**

Table A7.—School climate (math attitudes and academic expectations) predictors of student-level parameters of grade 4 math achievement—Continued

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.86	161.12	206	1422.89**
GENDER BETA COEFFICIENT	0.15	22.03	205	242.47*
RACE-ETHNICITY BETA COEFFICIENT	0.04	10.77	205	244.64†

[1]Average of five Gamma values. See technical notes for more information.

[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Parameter variance divided by total variance. Average of five reliability values.

[5]Average of five parameter variance values.

[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade Students, Restricted-Use Data Base.

Table A8.—Average within-school parameters of grade 4 geometry achievement

WITHIN-SCHOOL PARAMETERS	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)	213.13	1.09	195.28**
GENDER BETA COEFFICIENT	-0.14	1.05	-0.14
RACE-ETHNICITY BETA COEFFICIENT	-12.48	1.19	-10.76**
SES BETA COEFFICIENT	4.53	0.79	5.70**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.88	252.61	217	1570.89**
GENDER BETA COEFFICIENT	0.18	40.06	217	243.13
RACE-ETHNICITY BETA COEFFICIENT	0.07	23.66	217	240.76†

[1]Average of five Gamma values. See technical notes for more information.

[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Parameter variance divided by total variance. Average of five reliability values.

[5]Average of five parameter variance values.

[6]Average of five Chi-square tests.

NOTE: ** probability $\leq$ .01; * probability $\leq$ .05; † probability $\leq$ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade Students, Restricted-Use Data Base.

Table A9.—Student body characteristics predictors of student-level parameters of grade 4 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	214.64	1.01	213.12**
Percent African-American	-5.00	0.92	-5.42**
Percent Hispanic	-1.25	0.93	-1.35
Student body race-ethnicity unknown	6.14	6.41	0.96
Average SES	4.81	1.07	4.50**
GENDER BETA COEFFICIENT			
Intercept	-0.04	1.14	-0.04
Percent African-American	-0.25	0.95	-0.26
Percent Hispanic	0.07	1.04	0.06
Student body race-ethnicity unknown	-5.36	8.38	-0.64
Average SES	1.54	1.06	1.46
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-12.84	1.24	-10.37**
Percent African-American	-0.53	1.59	-0.33
Percent Hispanic	0.57	1.28	0.45
Student body race-ethnicity unknown	-5.38	7.83	-0.69
Average SES	0.56	1.64	0.34
SES BETA COEFFICIENT			
Intercept	4.51	0.79	5.71**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.84	185.33	213	1309.85**
GENDER BETA COEFFICIENT	0.17	37.28	213	240.29
RACE-ETHNICITY BETA COEFFICIENT	0.12	42.29	213	237.16†

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability $\leq$ .01; * probability $\leq$ .05; † probability $\leq$ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade Students, Restricted-Use Data Base.

Table A10a.—School resources predictors of student-level parameters of grade 4 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	214.96	1.03	209.32**
Percent African-American	-4.54	0.96	-4.72**
Percent Hispanic	-0.35	1.01	-0.34
Student body race-ethnicity unknown	8.61	7.51	1.15
Average SES	4.92	1.07	4.58**
School size (number of students)	-1.56	1.09	-1.44
Student/teacher ratio	-1.29	1.08	-1.19
Student/teacher ratio unknown	-7.81	9.82	-0.80
District instructional funds/student	-1.09	1.04	-1.05
District funds/student unknown	3.86	9.44	0.41
GENDER BETA COEFFICIENT			
Intercept	-0.10	1.17	-0.08
Percent African-American	-0.45	1.04	-0.43
Percent Hispanic	-0.18	1.10	-0.16
Student body race-ethnicity unknown	-4.96	9.28	-0.53
Average SES	1.48	1.10	1.34
School size (number of students)	0.96	1.49	0.65
Student/teacher ratio	-0 04	1.11	-0.04
Student/teacher ratio unknown	-1.97	9.78	-0.20
District instructional funds/student	0.33	1.25	0.27
District funds/student unknown	1.75	8.98	0.20
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-12.97	1.32	-9.82**
Percent African-American	0.01	1.66	0.01
Percent Hispanic	0.43	1.35	0.32
Student body race-ethnicity unknown	-7.76	9.60	-0.81
Average SES	0.92	1.59	0.58
School size (number of students)	0.02	1.46	0.01
Student/teacher ratio	1.31	1.33	0.99
Student/teacher ratio unknown	8.61	12.87	0.67
District instructional funds/student	-2.05	1.92	-1.07
District funds/student unknown	-5.74	11.87	-0.48
SES BETA COEFFICIENT			
Intercept	4.55	0.79	5.79**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.84	183.43	208	1290.52**
GENDER BETA COEFFICIENT	0.18	39.60	208	239.70
RACE-ETHNICITY BETA COEFFICIENT	0.13	46.09	208	232.85†

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability $\leq$ .01; * probability $\leq$ .05; † probability $\leq$ .10.

SOURCE: U. S. Department of Education, National Center for Educrtion Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 4th Grade Students, Restricted-Use Data Base.

Table A10b.—Detailed school resources predictors of student-level parameters of grade 4 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	215.23	1.12	192.59**
Percent African-American	-3.86	1.01	-3.82**
Percent Hispanic	-0.60	1.02	-0.58
Student body race-ethnicity unknown	8.01	7.46	1.07
Average SES	4.93	1.08	4.58**
School size (number of students)	-1.40	1.16	-1.20
Student/teacher ratio	-1.45	1.08	-1.34
Student/teacher ratio unknown	-7.74	9.74	-0.79
District instructional funds/student	-1.10	1.04	-1.05
District funds/student unknown	2.94	9.33	0.32
Computers per student	-0.51	0.97	-0.53
Computers per student unknown	3.98	3.85	1.03
Percent use computers as part of math instr.	2.52	1.10	2.28*
Percent using computers unknown	-10.20	4.55	-2.24*
GENDER BETA COEFFICIENT			
Intercept	-0.18	1.36	-0.13
Percent African-American	-0.59	1.13	-0.52
Percent Hispanic	-0.26	1.10	-0.24
Student body race-ethnicity unknown	-5.16	9.32	-0.55
Average SES	1.50	1.10	1.36
School size (number of students)	1.30	1.64	0.80
Student/teacher ratio	0.13	1.12	0.12
Student/teacher ratio unknown	-2.12	9.75	-0.22
District instructional funds/student	0.38	1.31	0.29
District funds/student unknown	1.91	8.93	0.21
Computers per student	0.98	1.05	0.93
Computers per student unknown	-1.00	4.31	-0.23
Percent use computers as part of math instr.	-0.64	1.22	-0.52
Percent using computers unknown	1.90	4.76	0.40
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-12.58	1.54	-8.19**
Percent African-American	-0.44	1.80	-0.24
Percent Hispanic	0.56	1.32	0.43
Student body race-ethnicity unknown	-8.32	10.00	-0.83
Average SES	1.35	1.59	0.85
School size (number of students)	0.38	1.57	0.24
Student/teacher ratio	1.62	1.32	1.23
Student/teacher ratio unknown	8.81	12.67	0.70
District instructional funds/student	-1.71	2.01	-0.85
District funds/student unknown	-5.60	11.84	-0.47
Computers per student	1.22	1.22	1.00
Computers per student unknown	-2.64	6.23	-0.42
Percent use computers as part of math instr.	-2.58	2.14	-1.20
Percent using computers unknown	1.86	6.28	0.30
SES BETA COEFFICIENT			
Intercept	4.53	0.78	5.78**

BEST COPY AVAILABLE

Table A10b.—Detailed school resources predictors of student-level parameters of grade 4 geometry achievement—Continued

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.83	177.98	204	1260.00**
GENDER BETA COEFFICIENT	0.19	40.68	204	236.67
RACE-ETHNICITY BETA COEFFICIENT	0.12	42.32	204	224.53

[1]Average of five Gamma values. See technical notes for more information.

[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Parameter variance divided by total variance. Average of five reliability values.

[5]Average of five parameter variance values.

[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP Math Assessment of 4th Grade Students, Restricted-Use Data Base.

Table A11.—Classroom instructional methods predictors of student-level parameters of grade 4 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	214.81	0.98	219.99**
Percent African-American	-4.68	1.00	-4.67**
Percent Hispanic	-0.75	0.94	-0.80
Student body race-ethnicity unknown	5.12	6.34	0.81
Average SES	4.60	1.08	4.26**
Work in small groups	-1.24	1.04	-1.18
Work with objects	2.02	1.10	1.83†
Do problems on worksheets	0.22	1.09	0.20
Do problems from textbook	2.96	1.04	2.83**
Take math tests	-1.33	1.10	-1.21
Use computer	0.96	1.02	0.94
Use calculator	-1.80	1.07	-1.68†
GENDER BETA COEFFICIENT			
Intercept	-0.01	1.16	-0.01
Percent African-American	-0.68	1.04	-0.65
Percent Hispanic	-0.44	1.17	-0.38
Student body race-ethnicity unknown	-5.19	8.30	-0.63
Average SES	1.17	1.14	1.03
Work in small groups	1.66	1.04	1.60
Work with objects	-0.23	1.56	-0.15
Do problems on worksheets	-1.18	1.32	-0.89
Do problems from textbook	-0.48	1.07	-0.45
Take math tests	-0.11	1.50	-0.07
Use computer	-1.17	1.27	-0.92
Use calculator	0.75	1.08	0.69
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-12.60	1.26	-10.01**
Percent African-American	0.11	1.69	0.07
Percent Hispanic	1.42	1.39	1.03
Student body race-ethnicity unknown	-4.97	8.22	-0.60
Average SES	0.88	1.72	0.51
Work in small groups	-0.30	1.47	-0.20
Work with objects	-0.64	1.71	-0.38
Do problems on worksheets	3.03	1.44	2.10*
Do problems from textbook	0.67	1.77	0.38
Take math tests	-0.21	1.48	-0.14
Use computer	-1.57	1.66	-0.95
Use calculator	-2.13	1.41	-1.51
SES BETA COEFFICIENT			
Intercept	4.54	0.79	5.72**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.83	171.23	206	1191.53**
GENDER BETA COEFFICIENT	0.18	38.97	206	235.92
RACE-ETHNICITY BETA COEFFICIENT	0.11	40.14	206	223.68

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 4th Grade Students, Restricted-Use Data Base.

Table A12.—School climate (math attitudes) predictors of student-level parameters of grade 4 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	214.81	1.00	214.36**
Percent African-American	-4.50	0.94	-4.78**
Percent Hispanic	-1.20	0.92	-1.31
Student body race-ethnicity unknown	4.52	6.46	0.70
Average SES	3.54	1.18	3.00**
Students feel math is useful	1.73	1.13	1.53
Students enjoy and feel competent in math	0.60	1.08	0.56
Students disagree that math is more for boys	1.95	1.10	1.77†
GENDER BETA COEFFICIENT			
Intercept	-0.03	1.15	-0.02
Percent African-American	-0.14	0.97	-0.15
Percent Hispanic	0.00	1.07	0.00
Student body race-ethnicity unknown	-4.43	8.59	-0.52
Average SES	1.22	1.13	1.08
Students feel math is useful	1.08	1.36	0.80
Students enjoy and feel competent in math	0.52	1.09	0.47
Students disagree that math is more for boys	-0.33	1.34	-0.25
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-12.81	1.26	-10.15**
Percent African-American	-0.74	1.61	-0.46
Percent Hispanic	0.59	1.29	0.46
Student body race-ethnicity unknown	-6.22	7.92	-0.79
Average SES	1.09	1.69	0.65
Students feel math is useful	-1.25	1.66	-0.75
Students enjoy and feel competent in math	-0.91	1.60	-0.57
Students disagree that math is more for boys	-0.05	1.52	-0.03
SES BETA COEFFICIENT			
Intercept	4.53	0.79	5.71**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.84	181.91	210	1253.76**
GENDER BETA COEFFICIENT	0.18	38.14	210	237.08
RACE-ETHNICITY BETA COEFFICIENT	0.13	47.50	210	233.60

[1] Average of five Gamma values. See technical notes for more information.
[2] Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3] Gamma divided by standard error. Probabilities based on a two-tailed test.
[4] Parameter variance divided by total variance. Average of five reliability values.
[5] Average of five parameter variance values.
[6] Average of five Chi-square tests.

NOTE: ** probability $\leq$ .01; * probability $\leq$ .05; † probability $\leq$ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP Math Assessment of 4th Grade Students, Restricted-Use Data Base.

Table A13.—School climate (math attitudes and student behavior and safety) predictors of student-level parameters of grade 4 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	215.58	1.03	208.34**
Percent African-American	-4.06	0.95	-4.27**
Percent Hispanic	-1.05	0.92	-1.14
Student body race-ethnicity unknown	3.30	6.43	0.51
Average SES	3.38	1.19	2.85**
Students feel math is useful	1.78	1.13	1.58
Students enjoy and feel competent in math	0.89	1.08	0.82
Students disagree that math is more for boys	2.11	1.10	1.93†
Index of problems in the school	-0.78	0.98	-0.79
Index of problems unknown	-8.14	12.25	-0.66
Percent students enrolled all year	0.83	1.05	0.80
Percent enrolled unknown	-2.21	11.84	-0.19
GENDER BETA COEFFICIENT			
Intercept	0.10	1.28	0.08
Percent African-American	0.14	1.02	0.14
Percent Hispanic	0.13	1.07	0.12
Student body race-ethnicity unknown	-4.83	8.61	-0.56
Average SES	1.15	1.17	0.99
Students feel math is useful	1.10	1.37	0.80
Students enjoy and feel competent in math	0.69	1.15	0.60
Students disagree that math is more for boys	-0.35	1.34	-0.26
Index of problems in the school	-0.23	1.08	-0.21
Index of problems unknown	16.73	15.04	1.11
Percent students enrolled all year	0.87	1.37	0.64
Percent enrolled unknown	-16.61	13.57	-1.22
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-12.61	1.33	-9.49**
Percent African-American	-0.47	1.67	-0.28
Percent Hispanic	0.64	1.28	0.50
Student body race-ethnicity unknown	-6.63	8.01	-0.83
Average SES	0.66	1.69	0.39
Students feel math is useful	-1.22	1.64	-0.75
Students enjoy and feel competent in math	-0.83	1.59	-0.52
Students disagree that math is more for boys	-0.03	1.54	-0.02
Index of problems in the school	-1.57	1.58	-0.99
Index of problems unknown	-1.46	30.06	-0.05
Percent students enrolled all year	-0.60	1.68	-0.36
Percent enrolled unknown	1.18	29.47	0.04
SES BETA COEFFICIENT			
Intercept	4.52	0.79	5.68**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.83	178.10	206	1235.18**
GENDER BETA COEFFICIENT	0.17	37.07	206	233.32
RACE-ETHNICITY BETA COEFFICIENT	0.13	49.86	206	231.21†

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade Students, Restricted-Use Data Base.

Table A14.—School climate (math attitudes and academic expectations) predictors of student-level parameters of grade 4 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	215.72	1.06	204.45**
Percent African-American	-3.99	0.96	-4.15**
Percent Hispanic	-1.12	0.92	-1.22
Student body race-ethnicity unknown	3.94	6.47	0.61
Average SES	3.64	1.18	3.09**
Students feel math is useful	1.90	1.13	1.67†
Students enjoy and feel competent in math	0.81	1.07	0.76
Students disagree that math is more for boys	2.17	1.10	1.97*
Amount of instruction in math	-0.54	1.08	-0.50
Amount of instruction unknown	-5.19	5.99	-0.87
Math identified as a special priority	-0.03	1.08	-0.03
Math special priority unknown	-5.59	5.75	-0.97
GENDER BETA COEFFICIENT			
Intercept	-0.07	1.35	-0.06
Percent African-American	-0.60	1.25	-0.48
Percent Hispanic	-0.01	1.05	-0.01
Student body race-ethnicity unknown	-4.68	9.00	-0.52
Average SES	1.59	1.28	1.24
Students feel math is useful	1.28	1.39	0.92
Students enjoy and feel competent in math	0.56	1.16	0.48
Students disagree that math is more for boys	-0.10	1.34	-0.08
Amount of instruction in math	-0.07	1.25	-0.06
Amount of instruction unknown	-4.09	5.53	-0.74
Math identified as a special priority	-1.13	1.11	-1.01
Math special priority unknown	4.35	5.91	0.74
Mean composite math score of grade sample	-1.19	1.63	-0.73
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-12.73	1.35	-9.40**
Percent African-American	-1.09	1.78	-0.61
Percent Hispanic	0.46	1.26	0.37
Student body race-ethnicity unknown	-8.35	7.76	-1.08
Average SES	1.04	1.92	0.54
Students feel math is useful	-1.45	1.75	-0.83
Students enjoy and feel competent in math	-0.50	1.65	-0.30
Students disagree that math is more for boys	0.17	1.61	0.11
Amount of instruction in math	2.60	1.52	1.71†
Amount of instruction unknown	1.02	8.27	0.12
Math identified as a special priority	-0.37	1.44	-0.26
Math special priority unknown	0.42	8.58	0.05
Mean composite math score of grade sample	-0.78	2.26	-0.35
SES BETA COEFFICIENT			
Intercept	4.49	0.79	5.67**

Table A14.—School climate (math attitudes and academic expectations) predictors of student-level parameters of grade 4 geometry achievement—Continued

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.83	176.76	206	1228.68**
GENDER BETA COEFFICIENT	0.18	37.64	205	232.74
RACE-ETHNICITY BETA COEFFICIENT	0.11	37.84	205	227.66

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability $\leq$ .01; * probability $\leq$ .05; † probability $\leq$ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 4th Grade Students, Restricted-Use Data Base.

Table A15.—Average within-school parameters of grade 8 math achievement

WITHIN-SCHOOL PARAMETERS	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)	260.28	1.36	191.70**
GENDER BETA COEFFICIENT	-0.63	0.84	-0.75
RACE-ETHNICITY BETA COEFFICIENT	-15.48	1.30	-11.96**
SES BETA COEFFICIENT	10.65	0.81	13.23**
TAKING ALGEBRA BETA COEFFICIENT	31.66	1.33	23.81**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.94	281.45	147	2225.69**
GENDER BETA COEFFICIENT	0.21	22.33	147	181.45*
RACE-ETHNICITY BETA COEFFICIENT	0.19	44.92	147	207.15**

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability $\leq$ .01; * probability $\leq$ .05; † probability $\leq$ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 8th Grade Students, Restricted-Use Data Base.

Table A16.—Student body characteristics predictors of student-level parameters of grade 8 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	262.96	0.82	320.01**
Percent African-American	-7.11	0.78	-9.08**
Percent Hispanic	-1.19	0.68	-1.76†
Average SES	9.31	0.88	10.55**
GENDER BETA COEFFICIENT			
Intercept	-0.60	0.84	-0.71
Percent African-American	-1.40	0.89	-1.57
Percent Hispanic	-1.12	0.85	-1.32
Average SES	-1.51	1.02	-1.48
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-16.32	1.41	-11.59**
Percent African-American	0.13	1.67	0.08
Percent Hispanic	1.78	1.16	1.53
Average SES	-1.67	1.44	-1.16
SES BETA COEFFICIENT			
Intercept	10.75	0.81	13.33**
TAKING ALGEBRA BETA COEFFICIENT			
Intercept	31.67	1.34	23.68**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.80	73.84	144	611.13**
GENDER BETA COEFFICIENT	0.20	21.92	144	177.81*
RACE-ETHNICITY BETA COEFFICIENT	0.16	35.96	144	190.72**

[1]Average of five Gamma values. See technical notes for more information.

[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Parameter variance divided by total variance. Average of five reliability values.

[5]Average of five parameter variance values.

[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 8th Grade Students, Restricted-Use Data Base.

Table A17.—School resources predictors of student-level parameters of grade 8 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	262.89	0.91	290.28**
Percent African-American	-7.18	0.80	-9.00**
Percent Hispanic	-1.38	0.76	-1.82†
Average SES	9.22	0.96	9.58**
School size (number of students)	0.04	0.84	0.04
Student/teacher ratio	0.59	0.88	0.67
Student/teacher ratio unknown	2.63	11.57	0.23
District instructional funds/student	0.39	0.93	0.41
District funds/student unknown	-2.91	11.27	-0.26
GENDER BETA COEFFICIENT			
Intercept	-0.50	0.92	-0.54
Percent African-American	-1.30	0.91	-1.43
Percent Hispanic	-0.82	0.93	-0.88
Average SES	-1.23	1.17	-1.05
School size (number of students)	-0.46	1.11	-0.41
Student/teacher ratio	-0.42	1.03	-0.41
Student/teacher ratio unknown	-11.85	15.61	-0.76
District instructional funds/student	-0.56	1.08	-0.52
District funds/student unknown	12.03	15.18	0.79
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-16.26	1.59	-10.26**
Percent African-American	0.27	1.69	0.16
Percent Hispanic	1.61	1.23	1.31
Average SES	-1.70	1.55	-1.09
School size (number of students)	-0.56	1.52	-0.37
Student/teacher ratio	0.81	1.47	0.55
Student/teacher ratio unknown	12.26	16.69	0.73
District instructional funds/student	0.73	1.50	0.49
District funds/student unknown	-11.46	16.43	-0.70
SES BETA COEFFICIENT			
Intercept	10.74	0.81	13.20**
TAKING ALGEBRA BETA COEFFICIENT			
Intercept	31.63	1.34	23.54**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.80	76.17	139	613.47**
GENDER BETA COEFFICIENT	0.21	22.65	139	177.05*
RACE-ETHNICITY BETA COEFFICIENT	0.17	39.18	139	185.73**

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 8th Grade Students, Restricted-Use Data Base.

Table A18.—Classroom instructional methods predictors of student-level parameters of grade 8 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	262.78	0.80	328.26**
Percent African-American	-6.83	0.83	-8.23**
Percent Hispanic	-0.78	0.68	-1.15
Average SES	9.21	0.86	10.66**
Work in small groups	-1.02	1.00	-1.02
Work with objects	0.64	0.86	0.75
Do problems on worksheets	1.70	1.08	1.58
Do problems from textbook	2.48	0.89	2.79**
Take math tests	-1.35	0.95	-1.42
Use computer	-2.16	0.87	-2.49*
Use calculator	1.26	0.93	1.35
GENDER BETA COEFFICIENT			
Intercept	-0.60	0.88	-0.68
Percent African-American	-0.84	1.00	-0.84
Percent Hispanic	-1.32	0.91	-1.45
Average SES	-1.58	1.03	-1.54
Work in small groups	-0.55	1.35	-0.40
Work with objects	-0.23	0.99	-0.24
Do problems on worksheets	-1.38	1.40	-0.98
Do problems from textbook	-1.20	1.05	-1.14
Take math tests	-1.07	1.35	-0.79
Use computer	0.83	1.16	0.72
Use calculator	0.67	1.52	0.44
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-16.00	1.42	-11.25**
Percent African-American	1.19	1.84	0.65
Percent Hispanic	1.49	1.20	1.24
Average SES	-1.81	1.44	-1.26
Work in small groups	2.53	1.72	1.47
Work with objects	3.16	1.42	2.23*
Do problems on worksheets	-0.15	2.00	-0.07
Do problems from textbook	-1.37	1.45	-0.95
Take math tests	-1.10	1.85	-0.60
Use computer	-0.74	1.66	-0.44
Use calculator	-1.27	1.53	-0.83
SES BETA COEFFICIENT			
Intercept	10.72	0.81	13.28**
TAKING ALGEBRA BETA COEFFICIENT			
Intercept	31.58	1.34	23.61**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.78	65.99	137	556.11**
GENDER BETA COEFFICIENT	0.22	24.36	137	176.91*
RACE-ETHNICITY BETA COEFFICIENT	0.14	31.52	137	178.91*

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability $\leq$ .01; * probability $\leq$ .05; † probability $\leq$ .10.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 8th Grade Students, Restricted-Use Data Base.

Table A19.—School climate (math attitudes) predictors of student-level parameters of grade 8 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	263.09	0.81	325.41**
Percent African-American	-7.43	0.77	-9.62**
Percent Hispanic	-1.13	0.68	-1.66†
Average SES	9.14	0.87	10.57**
Students feel math is useful	-1.25	0.93	-1.35
Students enjoy and feel competent in math	2.52	1.02	2.47*
Students disagree that math is more for boys	1.49	0.89	1.67†
GENDER BETA COEFFICIENT			
Intercept	-0.60	0.86	-0.70
Percent African-American	-1.30	0.91	-1.43
Percent Hispanic	-1.18	0.86	-1.38
Average SES	-1.52	1.04	-1.47
Students feel math is useful	0.12	1.13	0.10
Students enjoy and feel competent in math	-1.02	1.26	-0.81
Students disagree that math is more for boys	0.28	1.07	0.26
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-16.21	1.46	-11.10**
Percent African-American	0.07	1.71	0.04
Percent Hispanic	1.98	1.19	1.66†
Average SES	-1.50	1.45	-1.04
Students feel math is useful	0.32	1.54	0.21
Students enjoy and feel competent in math	2.81	1.88	1.50
Students disagree that math is more for boys	1.25	1.64	0.76
SES BETA COEFFICIENT			
Intercept	10.72	0.81	13.24**
TAKING ALGEBRA BETA COEFFICIENT			
Intercept	31.71	1.34	23.65**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.79	69.11	141	563.68**
GENDER BETA COEFFICIENT	0.21	22.66	141	176.21*
RACE-ETHNICITY BETA COEFFICIENT	0.15	34.12	141	187.09**

[1]Average of five Gamma values. See technical notes for more information.

[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Parameter variance divided by total variance. Average of five reliability values.

[5]Average of five parameter variance values.

[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education. National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 8th Grade Students, Restricted-Use Data Base.

Table A20.—School climate (math attitudes and student behavior and safety) predictors of student-level parameters of grade 8 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	263.06	0.83	316.69**
Percent African-American	-7.75	0.84	-9.22**
Percent Hispanic	-1.24	0.70	-1.77†
Average SES	9.26	0.89	10.44**
Students feel math is useful	-1.25	0.98	-1.27
Students enjoy and feel competent in math	2.59	1.08	2.40*
Students disagree that math is more for boys	1.68	0.93	1.81†
Absenteeism in grade	-0.58	0.87	-0.66
Students feel classes often disrupted	-0.58	1.01	-0.57
Students feel unsafe at school	1.35	1.12	1.20
GENDER BETA COEFFICIENT			
Intercept	-0.70	0.89	-0.79
Percent African-American	-1.33	0.98	-1.35
Percent Hispanic	-1.00	0.89	-1.13
Average SES	-1.39	1.09	-1.28
Students feel math is useful	-0.16	1.23	-0.13
Students enjoy and feel competent in math	-0.75	1.25	-0.60
Students disagree that math is more for boys	0.27	1.14	0.23
Absenteeism in grade	-0.12	1.15	-0.10
Students feel classes often disrupted	1.20	1.24	0.96
Students feel unsafe at school	-0.44	1.32	-0.34
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-16.75	1.50	-11.13**
Percent African-American	-0.52	1.83	-0.29
Percent Hispanic	1.71	1.24	1.38
Average SES	-1.18	1.47	-0.80
Students feel math is useful	0.26	1.68	0.15
Students enjoy and feel competent in math	3.25	1.94	1.68†
Students disagree that math is more for boys	1.80	1.69	1.06
Absenteeism in grade	1.40	1.69	0.83
Students feel classes often disrupted	0.95	1.91	0.50
Students feel unsafe at school	1.35	1.97	0.69
SES BETA COEFFICIENT			
Intercept	10.75	0.81	13.21**
TAKING ALGEBRA BETA COEFFICIENT			
Intercept	31.75	1.34	23.66**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.79	69.32	138	548.78**
GENDER BETA COEFFICIENT	0.21	22.53	138	177.43*
RACE-ETHNICITY BETA COEFFICIENT	0.15	33.14	138	180.30**

[1]Average of five Gamma values. See technical notes for more information.

[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Parameter variance divided by total variance. Average of five reliability values.

[5]Average of five parameter variance values.

[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 8th Grade Students, Restricted-Use Data Base.

Table A21.—School climate (math attitudes and academic expectations) predictors of student-level parameters of grade 8 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	262.82	0.78	336.36**
Percent African-American	-7.49	0.75	-10.05**
Percent Hispanic	-1.75	0.68	-2.57*
Average SES	7.86	0.90	8.73**
Students feel math is useful	-0.69	0.91	-0.76
Students enjoy and feel competent in math	2.34	0.99	2.37*
Students disagree that math is more for boys	1.95	0.87	2.24*
Percent of 8th grade students taking algebra	3.07	0.83	3.72**
GENDER BETA COEFFICIENT			
Intercept	-0.47	0.86	-0.55
Percent African-American	-1.45	1.44	-1.01
Percent Hispanic	-0.95	0.90	-1.06
Average SES	-0.78	1.54	-0.50
Students feel math is useful	-0.18	1.14	-0.16
Students enjoy and feel competent in math	-0.90	1.39	-0.65
Students disagree that math is more for boys	0.12	1.15	0.10
Percent of 8th grade students taking algebra	-1.30	1.08	-1.20
Mean composite math score of grade sample	-0.40	2.51	-0.16
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-15.77	1.50	-10.50**
Percent African-American	-0.61	2.07	-0.30
Percent Hispanic	2.55	1.27	2.00*
Average SES	0.69	2.16	0.32
Students feel math is useful	-0.54	1.59	-0.34
Students enjoy and feel competent in math	3.35	1.94	1.73†
Students disagree that math is more for boys	1.00	1.74	0.57
Percent of 8th grade students taking algebra	-3.02	1.60	-1.89†
Mean composite math score of grade sample	-1.47	3.09	-0.47
SES BETA COEFFICIENT			
Intercept	10.68	0.81	13.20**
TAKING ALGEBRA BETA COEFFICIENT			
Intercept	31.56	1.34	23.62**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.77	62.19	140	500.28**
GENDER BETA COEFFICIENT	0.20	21.29	139	175.23*
RACE-ETHNICITY BETA COEFFICIENT	0.15	33.83	139	183.58**

[1]Average of five Gamma values. See technical notes for more information.

[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Parameter variance divided by total variance. Average of five reliability values.

[5]Average of five parameter variance values.

[6]Average of five Chi-square tests.

NOTE: ** probability $\leq$ .01; * probability $\leq$ .05; † probability $\leq$ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 8th Grade Students, Restricted-Use Data Base.

Table A22.—Average within-school parameters of grade 8 geometry achievement

WITHIN-SCHOOL PARAMETERS	Gamma coefficient[1]	Standard error[2]	t Value[3]	
INTERCEPT (AVG. ACHIEVEMENT)	258.04	1.35	191.81**	
GENDER BETA COEFFICIENT	-1.23	1.03	-1.20	
RACE-ETHNICITY BETA COEFFICIENT	-14.00	1.68	-8.35**	
SES BETA COEFFICIENT	8.93	1.17	7.62**	
TAKING ALGEBRA BETA COEFFICIENT	27.84	1.47	18.91**	
Random within-school parameters	**Reliability[4]**	**Parameter variance (Tau)[5]**	**Degrees of freedom**	**Chi-square test of Tau > 0[6]**
INTERCEPT (AVG. ACHIEVEMENT)	0.90	260.16	147	1373.94**
GENDER BETA COEFFICIENT	0.17	26.51	147	161.90
RACE-ETHNICITY BETA COEFFICIENT	0.17	55.30	147	183.10*

[1]Average of five Gamma values. See technical notes for more information.

[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Parameter variance divided by total variance. Average of five reliability values.

[5]Average of five parameter variance values.

[6]Average of five Chi-square tests.

NOTE: ** probability $\leq$ .01; * probability $\leq$ .05; † probability $\leq$ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 8th Grade Students, Restricted-Use Data Base.

Table A23.—Student body characteristics predictors of student-level parameters of grade 8 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	260.42	0.89	294.06**
Percent African-American	-7.56	0.91	-8.33**
Percent Hispanic	-0.82	0.77	-1.06
Average SES	8.25	1.01	8.21**
GENDER BETA COEFFICIENT			
Intercept	-1.18	1.04	-1.14
Percent African-American	-1.20	1.09	-1.10
Percent Hispanic	-0.62	1.06	-0.58
Average SES	-0.03	1.21	-0.02
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.54	1.93	-7.53**
Percent African-American	-1.30	2.31	-0.56
Percent Hispanic	1.58	1.48	1.06
Average SES	-2.07	2.00	-1.04
SES BETA COEFFICIENT			
Intercept	9.03	1.17	7.72**
TAKING ALGEBRA BETA COEFFICIENT			
Intercept	27.82	1.49	18.68**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.74	77.77	144	449.88**
GENDER BETA COEFFICIENT	0.16	29.48	144	159.06
RACE-ETHNICITY BETA COEFFICIENT	0.16	51.58	144	171.32†

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 8th Grade Students, Restricted-Use Data Base.

Table A24.—School resources predictors of student-level parameters of grade 8 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	260.20	1.00	261.18**
Percent African-American	-7.73	0.91	-8.49**
Percent Hispanic	-1.21	0.86	-1.40
Average SES	7.88	1.15	6.87**
School size (number of students)	0.31	0.93	0.33
Student/teacher ratio	0.51	0.97	0.53
Student/teacher ratio unknown	4.83	13.08	0.37
District instructional funds/student	0.98	1.05	0.93
District funds/student unknown	-4.26	12.74	-0.33
GENDER BETA COEFFICIENT			
Intercept	-1.36	1.20	-1.13
Percent African-American	-1.25	1.13	-1.11
Percent Hispanic	-0.80	1.20	-0.67
Average SES	-0.30	1.53	-0.20
School size (number of students)	-0.11	1.29	-0.09
Student/teacher ratio	0.02	1.35	0.02
Student/teacher ratio unknown	-15.54	19.31	-0.80
District instructional funds/student	0.19	1.62	0.12
District funds/student unknown	17.98	18.74	0.96
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.26	2.24	-6.36**
Percent African-American	-1.09	2.35	-0.46
Percent Hispanic	1.48	1.48	1.00
Average SES	-2.06	2.02	-1.02
School size (number of students)	-1.03	2.30	-0.45
Student/teacher ratio	0 11	1.92	0.06
Student/teacher ratio unknown	17.84	21.34	0.84
District instructional funds/student	1.33	1.96	0.68
District funds/student unknown	-17.03	19.95	-0.85
SES BETA COEFFICIENT			
Intercept	9.00	1.19	7.59**
TAKING ALGEBRA BETA COEFFICIENT			
Intercept	27.82	1.50	18.57**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.74	79.09	139	456.77**
GENDER BETA COEFFICIENT	0.18	28.33	139	155.94
RACE-ETHNICITY BETA COEFFICIENT	0.15	50.16	139	165.29†

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 8th Grade Students, Restricted-Use Data Base.

BEST COPY AVAILABLE

Table A25.—Classroom instructional methods predictors of student-level parameters of grade 8 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	260.24	0.86	301.55**
Percent African-American	-7.09	0.94	-7.58**
Percent Hispanic	-0.45	0.81	-0.56
Average SES	8.11	0.98	8.28**
Work in small groups	-1.00	1.14	-0.87
Work with objects	1.73	0.92	1.88†
Do problems on worksheets	2.12	1.29	1.64
Do problems from textbook	2.07	0.97	2.12*
Take math tests	-2.01	1.08	-1.86†
Use computer	-2.03	0.99	-2.05*
Use calculator	1.47	1.01	1.45
GENDER BETA COEFFICIENT			
Intercept	-1.26	1.10	-1.14
Percent African-American	-1.04	1.33	-0.78
Percent Hispanic	-0.83	1.24	-0.67
Average SES	-0.12	1.23	-0.10
Work in small groups	-0.94	1.79	-0.52
Work with objects	-0.03	1.28	-0.02
Do problems on worksheets	-0.66	1.97	-0.33
Do problems from textbook	-1.32	1.36	-0.98
Take math tests	-0.27	1.64	-0.16
Use computer	1.10	1.46	0.76
Use calculator	0.79	2.36	0.33
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.01	1.98	-7.08**
Percent African-American	0.12	2.27	0.05
Percent Hispanic	1.29	1.54	0.84
Average SES	-2.09	1.92	-1.09
Work in small groups	2.36	2.15	1.10
Work with objects	3.24	1.76	1.84†
Do problems on worksheets	-0.93	2.82	-0.33
Do problems from textbook	-0.84	1.84	-0.46
Take math tests	-1.54	2.38	-0.65
Use computer	-0.01	2.15	0.00
Use calculator	-1.43	1.86	-0.77
SES BETA COEFFICIENT			
Intercept	9.01	1.16	7.79**
TAKING ALGEBRA BETA COEFFICIENT			
Intercept	27.78	1.49	18.69**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.71	66.79	137	407.25**
GENDER BETA COEFFICIENT	0.18	28.96	137	155.95
RACE-ETHNICITY BETA COEFFICIENT	0.15	46.93	137	163.21

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05. † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 8th Grade Students, Restricted-Use Data Base.

Table A26.—School climate (math attitudes) predictors of student-level parameters of grade 8 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	260.55	0.89	293.02**
Percent African-American	-7.76	0.91	-8.49**
Percent Hispanic	-0.79	0.78	-1.01
Average SES	8.11	1.01	8.06**
Students feel math is useful	-1.31	1.01	-1.30
Students enjoy and feel competent in math	1.50	1.13	1.32
Students disagree that math is more for boys	1.28	0.95	1.34
GENDER BETA COEFFICIENT			
Intercept	-1.18	1.08	-1.09
Percent African-American	-1.19	1.13	-1.05
Percent Hispanic	-0.72	1.07	-0.67
Average SES	-0.13	1.28	-0.10
Students feel math is useful	0.43	1.42	0.30
Students enjoy and feel competent in math	-0.62	1.43	-0.43
Students disagree that math is more for boys	0.57	1.49	0.38
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.54	1.93	-7.52**
Percent African-American	-1.45	2.23	-0.65
Percent Hispanic	1.92	1.54	1.25
Average SES	-1.84	1.94	-0.95
Students feel math is useful	-0.36	2.13	-0.17
Students enjoy and feel competent in math	4.10	2.61	1.57
Students disagree that math is more for boys	1.25	2.00	0.63
SES BETA COEFFICIENT			
Intercept	9.01	1.17	7.67**
TAKING ALGEBRA BETA COEFFICIENT			
Intercept	27.89	1.50	18.63**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.73	76.14	141	433.76**
GENDER BETA COEFFICIENT	0.18	29.19	141	160.14
RACE-ETHNICITY BETA COEFFICIENT	0.14	44.16	141	166.33†

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 8th Grade Students, Restricted-Use Data Base.

Table A27.—School climate (math attitudes and student behavior and safety) predictors of student-level parameters of grade 8 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	260.50	0.90	288.44**
Percent African-American	8.06	1.00	-8.04**
Percent Hispanic	-0.89	0.81	-1.10
Average SES	8.22	1.03	7.96**
Students feel math is useful	-1.35	1.07	-1.26
Students enjoy and feel competent in math	1.61	1.19	1.35
Students disagree that math is more for boys	1.45	1.01	1.43
Absenteeism in grade	-0.45	0.97	-0.47
Students feel classes often disrupted	-0.37	1.14	-0.32
Students feel unsafe at school	1.16	1.22	0.95
GENDER BETA COEFFICIENT			
Intercept	-1.30	1.12	-1.16
Percent African-American	-1.34	1.19	-1.13
Percent Hispanic	-0.77	1.12	-0.68
Average SES	0.01	1.37	0.01
Students feel math is useful	0.45	1.56	0.29
Students enjoy and feel competent in math	-0.45	1.48	-0.30
Students disagree that math is more for boys	0.74	1.49	0.49
Absenteeism in grade	0.51	1.61	0.32
Students feel classes often disrupted	0.40	1.47	0.27
Students feel unsafe at school	0.29	1.54	0.19
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.96	2.06	-7.27**
Percent African-American	-1.79	2.26	-0.79
Percent Hispanic	1.64	1.55	1.06
Average SES	-1.71	1.94	-0.88
Students feel math is useful	-0.18	2.28	-0.08
Students enjoy and feel competent in math	4.33	2.73	1.59
Students disagree that math is more for boys	1.61	2.08	0.77
Absenteeism in grade	1.43	2.12	0.68
Students feel classes often disrupted	0.19	2.26	0.08
Students feel unsafe at school	1.00	2.45	0.41
SES BETA COEFFICIENT			
Intercept	9.04	1.17	7.74**
TAKING ALGEBRA BETA COEFFICIENT			
Intercept	27.92	1.50	18.65**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.74	76.91	138	426.46**
GENDER BETA COEFFICIENT	0.16	25.21	138	159.14
RACE-ETHNICITY BETA COEFFICIENT	0.15	49.06	138	163.36

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 8th Grade Students, Restricted-Use Data Base.

Table A28.—School climate (math attitudes and academic expectations) predictors of student-level parameters of grade 8 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	260.33	0.88	295.22**
Percent African-American	-7.81	0.90	-8.69**
Percent Hispanic	-1.31	0.79	-1.65†
Average SES	7.03	1.07	6.59**
Students feel math is useful	-0.84	1.01	-0.83
Students enjoy and feel competent in math	1.36	1.12	1.22
Students disagree that math is more for boys	1.65	0.95	1.74†
Percent of 8th grade students taking algebra	2.54	0.94	2.71**
GENDER BETA COEFFICIENT			
Intercept	-1.01	1.10	-0.92
Percent African-American	-2.19	1.87	-1.17
Percent Hispanic	-0.81	1.09	-0.74
Average SES	1.27	1.77	0.72
Students feel math is useful	0.14	1.43	0.10
Students enjoy and feel competent in math	-0.31	1.50	-0.21
Students disagree that math is more for boys	0.69	1.59	0.43
Percent of 8th grade students taking algebra	-0.43	1.18	-0.37
Mean composite math score of grade sample	-2.44	3.11	-0.78
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.03	1.98	-7.10**
Percent African-American	-3.31	3.30	-1.00
Percent Hispanic	1.97	1.66	1.19
Average SES	1.08	3.33	0.33
Students feel math is useful	-1.06	2.36	-0.45
Students enjoy and feel competent in math	4.98	3.09	1.61
Students disagree that math is more for boys	1.61	2.12	0.76
Percent of 8th grade students taking algebra	-1.26	2.36	-0.54
Mean composite math score of grade sample	-4.52	5.77	-0.78
SES BETA COEFFICIENT			
Intercept	8.99	1.17	7.67*
TAKING ALGEBRA BETA COEFFICIENT			
Intercept	27.76	1.47	18.88**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.73	72.78	140	413.82**
GENDER BETA COEFFICIENT	0.15	23.11	139	159.05
RACE-ETHNICITY BETA COEFFICIENT	0.17	57.89	139	170.12†

[1]Average of five Gamma values. See technical notes for more information.

[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Parameter variance divided by total variance. Average of five reliability values.

[5]Average of five parameter variance values.

[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01, * probability ≤ .05, † probability ≤ .10.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP Math Assessment of 8th Grade Students, Restricted-Use Data Base.

Table A29.—Average within-school parameters of grade 12 math achievement

WITHIN-SCHOOL PARAMETERS	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)	291.98	1.37	213.67**
GENDER BETA COEFFICIENT	-2.53	0.97	-2.60**
RACE-ETHNICITY BETA COEFFICIENT	-14.13	1.70	-8.33**
SES BETA COEFFICIENT	12.36	0.85	14.48**
YEARS OF CALCULUS BETA COEFFICIENT	17.58	0.81	21.59**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.93	299.05	136	1529.51**
GENDER BETA COEFFICIENT	0.18	24.18	136	160.19†
RACE-ETHNICITY BETA COEFFICIENT	0.26	72.88	136	181.04**

[1] Average of five Gamma values. See technical notes for more information.

[2] Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3] Gamma divided by standard error. Probabilities based on a two-tailed test.

[4] Parameter variance divided by total variance. Average of five reliability values.

[5] Average of five parameter variance values

[6] Average of five Chi-square tests.

NOTE: ** probability $\leq$ .01; * probability $\leq$ .05; † probability $\leq$ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 12th Grade Students, Restricted-Use Data Base.

Table A30.—Student body characteristics predictors of student-level parameters of grade 12 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	292.80	0.83	352.69**
Percent African-American	-6.55	0.82	-8.03**
Percent Hispanic	0.91	0.71	1.27
Average SES	11.64	0.89	13.07**
GENDER BETA COEFFICIENT			
Intercept	-2.66	0.99	-2.70**
Percent African-American	-0.02	1.08	-0.02
Percent Hispanic	-0.55	0.95	-0.58
Average SES	0.30	1.18	0.26
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.04	1.89	-7.43**
Percent African-American	-0.97	2.00	-0.49
Percent Hispanic	0.35	1.39	0.25
Average SES	-0.08	1.93	-0.04
SES BETA COEFFICIENT			
Intercept	12.34	0.85	14.44**
YEARS OF CALCULUS BETA COEFFICIENT			
Intercept	17.59	0.82	21.45**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.77	76.66	133	574.49**
GENDER BETA COEFFICIENT	0.19	25.83	133	159.24†
RACE-ETHNICITY BETA COEFFICIENT	0.26	74.72	133	179.43**

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 12th Grade Students, Restricted-Use Data Base.

Table A31.—School resources predictors of student-level parameters of grade 12 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	292.55	0.89	328.72**
Percent African-American	-6.60	0.82	-8.04**
Percent Hispanic	0.48	0.79	0.61
Average SES	10.98	0.98	11.19**
School size (number of students)	1.10	0.81	1.35
Student/teacher ratio	-0.89	0.95	-0.94
Student/teacher ratio unknown	-0.71	3.64	-0.19
District instructional funds/student[4]	0.76	0.88	0.87
GENDER BETA COEFFICIENT			
Intercept	-2.39	1.04	-2.30*
Percent African-American	0.20	1.09	0.18
Percent Hispanic	0.11	1.04	0.11
Average SES	1.60	1.27	1.26
School size (number of students)	-1.96	1.09	-1.80†
Student/teacher ratio	1.65	1.23	1.34
Student/teacher ratio unknown	2.19	5.17	0.42
District instructional funds/student	-1.24	1.15	-1.08
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.06	2.08	-6.75**
Percent African-American	-0.50	1.91	-0.26
Percent Hispanic	-0.18	1.40	-0.13
Average SES	0.30	1.98	0.15
School size (number of students)	-3.18	1.65	-1.92†
Student/teacher ratio	4.57	1.91	2.40*
Student/teacher ratio unknown	14.82	7.27	2.04*
District instructional funds/student	4.17	1.73	2.41*
SES BETA COEFFICIENT			
Intercept	12.39	0.85	14.56**
YEARS OF CALCULUS BETA COEFFICIENT			
Intercept	17.51	0.81	21.50**

Random within-school parameters	Reliability[5]	Parameter variance (Tau)[6]	Degrees of freedom	Chi-square test of Tau > 0[7]
INTERCEPT (AVG. ACHIEVEMENT)	0.77	77.21	129	559.48**
GENDER BETA COEFFICIENT	0.18	23.84	129	150.96
RACE-ETHNICITY BETA COEFFICIENT	0.19	50.67	129	169.72**

[1]Average of five Gamma values. See technical notes for more information.

[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Although there were cases for which district instructional funds were unknown, these were the same cases for which student teacher ratio was unknown.

[5]Parameter variance divided by total variance. Average of five reliability values.

[6]Average of five parameter variance values.

[7]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 12th Grade Students, Restricted-Use Data Base.

Table A32.—Classroom instructional methods predictors of student-level parameters of grade 12 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	292.55	0.77	379.35**
Percent African-American	-6.31	0.85	-7.47**
Percent Hispanic	0.22	0.69	0.32
Average SES	8.60	0.97	8.85**
Work in small groups	0.91	0.92	-1.00
Work with objects	-1.47	0.90	-1.63
Do problems on worksheets	-0.75	0.96	-0.77
Do problems from textbook	3.03	1.31	2.32*
Take math tests	-0.63	1.20	-0.52
Use computer	-0.45	0.95	-0.48
Use calculator	3.53	1.10	3.22**
Write math proofs	1.99	0.95	2.11*
Formulate own math problems	-1.52	1.04	-1.47
GENDER BETA COEFFICIENT			
Intercept	-2.52	0.96	-2.62**
Percent African-American	0.65	1.21	0.54
Percent Hispanic	-0.21	1.00	-0.21
Average SES	1.86	1.29	1.44
Work in small groups	0.96	1.19	0.81
Work with objects	1.71	1.21	1.42
Do problems on worksheets	0.95	1.24	0.77
Do problems from textbook	0.16	1.68	0.09
Take math tests	-2.88	1.54	-1.87†
Use computer	-0.92	1.28	-0.72
Use calculator	-0.56	1.52	-0.37
Write math proofs	-0.71	1.36	-0.52
Formulate own math problems	1.37	1.32	1.04
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.88	2.00	-7.42**
Percent African-American	0.23	2.26	0.10
Percent Hispanic	0.78	1.46	0.53
Average SES	2.94	2.27	1.30
Work in small groups	0.58	2.41	0.24
Work with objects	1.04	2.11	0.49
Do problems on worksheets	2.05	2.35	0.87
Do problems from textbook	-4.36	3.53	-1.24
Take math tests	-2.65	3.17	-0.84
Use computer	-3.77	2.46	-1.53
Use calculator	-2.41	2.74	-0.88
Write math proofs	-0.20	2.10	-0.09
Formulate own math problems	1.53	2.61	0.58
SES BETA COEFFICIENT			
Intercept	12.26	0.85	14.45**
YEARS OF CALCULUS BETA COEFFICIENT			
Intercept	17.51	0.82	21.42**

Table A32.—Classroom instructional methods predictors of student-level parameters of grade 12 math achievement—Continued

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.72	58.93	124	439.61**
GENDER BETA COEFFICIENT	0.12	15.26	124	159.20*
RACE-ETHNICITY BETA COEFFICIENT	0.26	75.43	124	173.19**

[1]Average of five Gamma values. See technical notes for more information.

[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Parameter variance divided by total variance. Average of five reliability values.

[5]Average of five parameter variance values.

[6]Average of five Chi-square tests.

NOTE: ** probability $\leq$ .01; * probability $\leq$ .05; † probability $\leq$ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP, Math Assessment of 4th Grade Students, Restricted-Use Data Base.

Table A33.—School climate (math attitudes) predictors of student-level parameters of grade 12 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	292.59	0.74	394.17**
Percent African-American	-7.17	0.77	-9.30**
Percent Hispanic	1.03	0.65	1.57
Average SES	10.79	0.83	12.98**
Students feel math is useful	-3.33	0.97	-3.43**
Students enjoy and feel competent in math	6.66	0.95	7.01**
Students disagree that math is more for boys	-0.06	0.84	-0.07
GENDER BETA COEFFICIENT			
Intercept	-2.64	0.99	-2.67**
Percent African-American	-0.09	1.11	-0.08
Percent Hispanic	-0.58	0.96	-0.61
Average SES	0.54	1.20	0.45
Students feel math is useful	0.80	1.43	0.56
Students enjoy and feel competent in math	-1.07	1.30	-0.82
Students disagree that math is more for boys	-0.71	1.11	-0.65
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.17	1.90	-7.47**
Percent African-American	-1.53	2.07	-0.74
Percent Hispanic	0.40	1.45	0.28
Average SES	-0.48	2.06	-0.23
Students feel math is useful	-1.31	2.41	-0.54
Students enjoy and feel competent in math	0.00	2.27	0.00
Students disagree that math is more for boys	-3.13	1.89	-1.66†
SES BETA COEFFICIENT			
Intercept	12.35	0.85	14.49**
YEARS OF CALCULUS BETA COEFFICIENT			
Intercept	17.57	0.82	21.49**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.71	54.69	130	466.56**
GENDER BETA COEFFICIENT	0.18	24.70	130	156.30†
RACE-ETHNICITY BETA COEFFICIENT	0.28	83.95	130	174.06**

[1]Average of five Gamma values. See technical notes for more information.

[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Parameter variance divided by total variance. Average of five reliability values.

[5]Average of five parameter variance values.

[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP Math Assessment of 8th Grade Students, Restricted-Use Data Base.

Table A34.—School climate (math attitudes and student behavior and safety) predictors of student-level parameters of grade 12 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	293.06	0.71	412.02**
Percent African-American	-6.25	0.83	-7.56**
Percent Hispanic	1.03	0.65	1.59
Average SES	9.98	0.82	12.14**
Students feel math is useful	-3.22	0.93	-3.48**
Students enjoy and feel competent in math	6.42	0.91	7.07**
Students disagree that math is more for boys	-0.30	0.80	-0.37
Absenteeism in grade	0.84	0.78	1.07
Students feel classes often disrupted	-2.84	0.83	-3.43**
Students feel unsafe at school	-1.36	0.93	-1.46
GENDER BETA COEFFICIENT			
Intercept	-2.69	1.00	-2.69**
Percent African-American	-0.24	1.24	-0.19
Percent Hispanic	-0.21	1.00	-0.22
Average SES	0.55	1.24	0.44
Students feel math is useful	0.77	1.44	0.53
Students enjoy and feel competent in math	-1.06	1.30	-0.82
Students disagree that math is more for boys	-0.60	1.12	-0.54
Absenteeism in grade	-1.78	1.10	-1.62
Students feel classes often disrupted	2.19	1.19	1.84†
Students feel unsafe at school	-0.88	1.47	-0.60
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-12.80	1.96	-6.52**
Percent African-American	-0.80	2.18	-0.37
Percent Hispanic	0.37	1.46	0.25
Average SES	-1.11	2.09	-0.53
Students feel math is useful	-0.93	2.38	-0.39
Students enjoy and feel competent in math	-0.96	2.30	-0.41
Students disagree that math is more for boys	-3.37	1.84	-1.83†
Absenteeism in grade	-0.05	2.02	-0.03
Students feel classes often disrupted	-2.90	2.09	-1.39
Students feel unsafe at school	-1.56	2.27	-0.69
SES BETA COEFFICIENT			
Intercept	12.37	0.86	14.45**
YEARS OF CALCULUS BETA COEFFICIENT			
Intercept	17.54	0.82	21.31**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.67	45.73	127	381.53**
GENDER BETA COEFFICIENT	0.19	25.84	127	151.51†
RACE-ETHNICITY BETA COEFFICIENT	0.26	72.87	127	176.24**

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05, † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 12th Grade Students, Restricted-Use Data Base

Table A35.—School climate (math attitudes and academic expectations) predictors of student-level parameters of grade 12 math achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	292.38	0.69	423.26**
Percent African-American	-6.90	0.73	-9.49**
Percent Hispanic	0.43	0.63	0.68
Average SES	8.22	0.92	8.95**
Students feel math is useful	-1.90	0.94	-2.02*
Students enjoy and feel competent in math	4.92	0.94	5.26**
Students disagree that math is more for boys	-0.21	0.79	-0.26
Percent of students on academic/college prep	3.84	0.89	4.31**
Mean years 12th graders have taken calculus	2.12	0.76	2.80**
GENDER BETA COEFFICIENT			
Intercept	-2.48	1.01	-2.45*
Percent African-American	0.84	1.39	0.60
Percent Hispanic	-0.32	0.99	-0.32
Average SES	0.58	1.87	0.31
Students feel math is useful	0.56	1.53	0.37
Students enjoy and feel competent in math	-0.98	1.52	-0.64
Students disagree that math is more for boys	-0.68	1.12	-0.61
Percent of students on academic/college prep	-2.42	1.47	-1.65†
Mean years 12th graders have taken calculus	-1.43	1.15	-1.25
Mean composite math score of grade sample	2.90	2.22	1.31
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-13.98	1.90	-7.35**
Percent African-American	0.89	2.78	0.32
Percent Hispanic	0.18	1.47	0.12
Average SES	-2.93	2.76	-1.06
Students feel math is useful	-0.90	2.55	-0.35
Students enjoy and feel competent in math	-1.44	2.67	-0.54
Students disagree that math is more for boys	-3.63	1.88	-1.94†
Percent of students on academic/college prep	-3.36	2.39	-1.41
Mean years 12th graders have taken calculus	0.26	1.87	0.14
Mean composite math score of grade sample	6.16	4.11	1.50
SES BETA COEFFICIENT			
Intercept	12.36	0.85	14.48**
YEARS OF CALCULUS BETA COEFFICIENT			
Intercept	17.56	0.82	21.44**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.65	42.97	128	365.66**
GENDER BETA COEFFICIENT	0.22	30.15	127	161.71*
RACE-ETHNICITY BETA COEFFICIENT	0.29	85.40	127	181.36**

[1]Average of five Gamma values. See technical notes for more information.

[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Parameter variance divided by total variance. Average of five reliability values.

[5]Average of five parameter variance values.

[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 12th Grade Students, Restricted-Use Data Base.

BEST COPY AVAILABLE

Table A36.—-Average within-school parameters of grade 12 geometry achievement

WITHIN-SCHOOL PARAMETERS	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)	292.08	1.60	182.54**
GENDER BETA COEFFICIENT	-6.47	1.26	-5.12**
RACE-ETHNICITY BETA COEFFICIENT	-14.30	1.88	-7.63**
SES BETA COEFFICIENT	7.49	1.29	5.81**
YEARS OF GEOMETRY BETA COEFFICIENT	23.58	0.63	37.35**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.93	392.68	136	1563.91**
GENDER BETA COEFFICIENT	0.20	35.61	136	154.35
RACE-ETHNICITY BETA COEFFICIENT	0.25	90.50	136	188.07**

[1]Average of five Gamma values. See technical notes for more information

[2]Average of five standard error values plus standard error of the five Gammas See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Parameter variance divided by total variance. Average of five reliability values.

[5]Average of five parameter variance values.

[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10

SOURCE: U. S. Department of Education. National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 12th Grade Students, Restricted-Use Data Base.

Table A37.—Student body characteristics predictors of student-level parameters of grade 12 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	292.97	1.03	283.98**
Percent African-American	-7.18	0.97	-7.44**
Percent Hispanic	1.11	0.86	1.30
Average SES	13.18	1.11	11.90**
GENDER BETA COEFFICIENT			
Intercept	-6.67	1.35	-4.93**
Percent African-American	0.79	1.38	0.57
Percent Hispanic	-0.74	1.26	-0.59
Average SES	1.35	1.79	0.76
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.15	2.08	-6.81**
Percent African-American	-2.00	2.54	-0.79
Percent Hispanic	0.54	1.76	0.31
Average SES	-2.80	2.52	-1.11
SES BETA COEFFICIENT			
Intercept	7.45	1.31	5.69**
YEARS OF GEOMETRY BETA COEFFICIENT			
Intercept	23.59	0.63	37.57**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.79	112.78	133	627.64**
GENDER BETA COEFFICIENT	0.20	38.28	133	151.05†
RACE-ETHNICITY BETA COEFFICIENT	0.23	88.02	133	179.14**

[1] Average of five Gamma values. See technical notes for more information.

[2] Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3] Gamma divided by standard error. Probabilities based on a two-tailed test.

[4] Parameter variance divided by total variance. Average of five reliability values.

[5] Average of five parameter variance values.

[6] Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 12th Grade Students, Restricted-Use Data Base.

Table A38.—School resources predictors of student-level parameters of grade 12 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	292.62	1.11	263.22**
Percent African-American	-7.25	0.96	-7.53**
Percent Hispanic	0.56	0.94	0.59
Average SES	12.12	1 20	10.07**
School size (number of students)	1.46	0.95	1.53
Student/teacher ratio	-1.64	1.14	-1.44
Student/teacher ratio unknown	0.37	4.47	0.08
District instructional funds/student[4]	.24	1.03	1.21
GENDER BETA COEFFICIENT			
Intercept	-6.50	1.45	-4.49**
Percent African-American	0.86	1.38	0.62
Percent Hispanic	-0.22	1.43	-0.16
Average SES	2.37	1.90	1.25
School size (number of students)	-0.72	1.29	-0.56
Student/teacher ratio	0.33	1.44	0.23
Student/teacher ratio unknown	-0.25	6.71	-0.04
District instructional funds/student	-1.94	1.38	-1.41
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-13.63	2.40	-5.69**
Percent African-American	-1.59	2.51	-0.64
Percent Hispanic	0.43	1.80	0.24
Average SES	-2.56	2.70	-0.95
School size (number of students	-3.07	2.33	-1.32
Student/teacher ratio	3.03	2.32	1.31
Student/teacher ratio unknown	9.14	8.52	1.07
District instructional funds/student	3.11	2.09	1.49
SES BETA COEFFICIENT			
Intercept	7.51	1.29	5.84**
YEARS OF GEOMETRY BETA COEFFICIENT			
Intercept	23.50	0.63	37.60**

Random within-school parameters	Reliability[5]	Parameter variance (Tau)[6]	Degrees of freedom	Chi-square test of Tau > 0[7]
INTERCEPT (AVG. ACHIEVEMENT)	0.79	111.08	129	605.73**
GENDER BETA COEFFICIENT	0.19	33.42	129	150.27
RACE-ETHNICITY BETA COEFFICIENT	0.21	74.51	129	176.86**

[1]Average of five Gamma values. See technical notes for more information.

[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.

[3]Gamma divided by standard error. Probabilities based on a two-tailed test.

[4]Although there were cases for which district instructional funds were unknown, these were the same cases for which student teacher ratio was unknown.

[5]Parameter variance divided by total variance. Average of five reliability values.

[6]Average of five parameter variance values.

[7]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05, † probability ≤ .10.

SOURCE U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 12th Grade Students, Restricted-Use Data Base.

Table A39.—Classroom instructional methods predictors of student-level parameters of grade 12 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	292.71	0.97	303.32••
Percent African-American	-6.79	1.04	-6.52••
Percent Hispanic	0.31	0.83	0.37
Average SES	9.69	1.19	8.11••
Work in small groups	-0.53	1.14	-0.47
Work with objects	-2.00	1.07	-1.86†
Do problems on worksheets	-0.93	1.18	-0.78
Do problems from textbook	2.52	1.64	1.54
Take math tests	0.15	1.51	0.10
Use computer	-0.32	1.18	-0.28
Use calculator	4.27	1.35	3.16••
Write math proofs	2.32	1.19	1.95†
Formulate own math problems	-2.48	1.43	-1.74†
GENDER BETA COEFFICIENT			
Intercept	-6.91	1.32	-5.22••
Percent African-American	1.02	1.43	0.71
Percent Hispanic	-0.52	1.25	-0.41
Average SES	3.06	1.85	1.65†
Work in small groups	1.07	1.39	0.77
Work with objects	2.71	1.46	1.86†
Do problems on worksheets	0.90	1.44	0.63
Do problems from textbook	-1.23	2.01	-0.61
Take math tests	-1.63	1.81	-0.90
Use computer	-3.21	1.54	-2.08•
Use calculator	0.19	1.85	0.10
Write math proofs	-0.77	1.52	-0.50
Formulate own math problems	1.95	1.58	1.23
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-13.99	2.28	-6.15••
Percent African-American	0.64	2.58	0.25
Percent Hispanic	40.83	1.87	0.4
Average SES	-1.29	2.85	-0.45
Work in small groups	-0.52	3.21	-0.16
Work with objects	-0.15	3.34	-0.04
Do problems on worksheets	3.17	2.67	1.19
Do problems from textbook	-2.06	3.92	-0.53
Take math tests	-6.71	3.99	-1.68†
Use computer	0.95	3.22	0.30
Use calculator	-0.16	3.19	-0.05
Write math proofs	0.93	2.43	0.38
Formulate own math problems	-0.05	2.93	-0.02
SES BETA COEFFICIENT			
Intercept	7.46	1.30	5.74••
YEARS OF GEOMETRY BETA COEFFICIENT			
Intercept	23.59	0.63	37.59••

Table A39.—Classroom instructional methods predictors of student-level parameters of grade 12 geometry achievement —Continued

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.76	90.97	124	490.71**
GENDER BETA COEFFICIENT	0.14	22.33	124	150.49†
RACE-ETHNICITY BETA COEFFICIENT	0.24	86.97	124	167.13**

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability $\leq .01$; * probability $\leq .05$; † probability $\leq .10$.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP: Math Assessment of 4th Grade Students, Restricted-Use Data Base.

Table A40.—School climate (math attitudes) predictors of student-level parameters of grade 12 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	292.70	0.94	312.27**
Percent African-American	-7.78	0.93	-8.41**
Percent Hispanic	1.31	0.79	1.66†
Average SES	12.09	1.03	11.74**
Students feel math is useful	-4.49	1.16	-3.88**
Students enjoy and feel competent in math	7.58	1.22	6.21**
Students disagree that math is more for boys	-0.28	1.02	-0.28
GENDER BETA COEFFICIENT			
Intercept	-6.63	1.36	-4.87**
Percent African-American	0.81	1.41	0.58
Percent Hispanic	-0.84	1.26	-0.66
Average SES	1.89	1.80	1.05
Students feel math is useful	1.99	1.61	1.24
Students enjoy and feel competent in math	-2.65	1.60	-1.65†
Students disagree that math is more for boys	-0.84	1.29	-0.66
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-14.04	2.10	-6.69**
Percent African-American	-1.93	2.63	-0.73
Percent Hispanic	0.66	1.81	0.36
Average SES	-3.59	2.49	-1.44
Students feel math is useful	-1.87	2.64	-0.71
Students enjoy and feel competent in math	-2.83	2.57	-1.10
Students disagree that math is more for boys	-3.07	2.04	-1.51
SES BETA COEFFICIENT			
Intercept	7.43	1.31	5.69**
YEARS OF GEOMETRY BETA COEFFICIENT			
Intercept	23.62	0.63	37.74**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.74	82.99	130	514.15**
GENDER BETA COEFFICIENT	0.19	34.05	130	140.88
RACE-ETHNICITY BETA COEFFICIENT	0.23	83.02	130	173.72**

[1]Average of five Gamma values See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP, Math Assessment of 8th Grade Students, Restricted-Use Data Base.

Table A41.—School climate (math attitudes and student behavior and safety) predictors of student-level parameters of grade 12 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	293.21	0.90	327.35**
Percent African-American	-6.78	0.98	-6.94**
Percent Hispanic	1.26	0.78	1.61
Average SES	11.28	1.03	10.91**
Students feel math is useful	-4.38	1.11	-3.94**
Students enjoy and feel competent in math	7.32	1.19	6.17**
Students disagree that math is more for boys	-0.55	0.98	-0.56
Absenteeism in grade	1.19	0.94	1.27
Students feel classes often disrupted	-3.22	1.01	-3.17**
Students feel unsafe at school	-1.33	1.12	-1.19
GENDER BETA COEFFICIENT			
Intercept	6.80	1.37	-4.98**
Percent African-American	0.30	1.56	0.19
Percent Hispanic	-0.65	1.23	-0.53
Average SES	2.04	1.88	1.08
Students feel math is useful	1.82	1.60	1.14
Students enjoy and feel competent in math	-2.60	1.58	-1.64
Students disagree that math is more for boys	-0.73	1.27	-0.57
Absenteeism in grade	-1.96	1.43	-1.37
Students feel classes often disrupted	1.65	1.63	1.02
Students feel unsafe at school	0.28	1.72	0.16
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-12.92	2.26	-5.73**
Percent African-American	-1.72	2.83	-0.61
Percent Hispanic	0.88	1.86	0.47
Average SES	-4.45	2.45	-1.82†
Students feel math is useful	-1.54	2.61	-0.59
Students enjoy and feel competent in math	-3.63	2.67	-1.36
Students disagree that math is more for boys	-2.98	2.03	-1.47
Absenteeism in grade	2.78	2.39	-1.16
Students feel classes often disrupted	-1.59	3.34	-0.47
Students feel unsafe at school	-1.45	2.71	-0.53
SES BETA COEFFICIENT			
Intercept	7.44	1.31	5.67**
YEARS OF GEOMETRY BETA COEFFICIENT			
Intercept	23.61	0.62	38.00**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.71	71.14	127	427.74**
GENDER BETA COEFFICIENT	0.16	27.56	127	138.32
RACE-ETHNICITY BETA COEFFICIENT	0.21	73.51	127	170.24**

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance Average of five reliability values
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE. U S Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 12th Grade Students, Restricted-Use Data Base.

Table A42.—School climate (math attitudes and academic expectations) predictors of student-level parameters of grade 12 geometry achievement

WITHIN-SCHOOL PARAMETERS Between-school predictors	Gamma coefficient[1]	Standard error[2]	t Value[3]
INTERCEPT (AVG. ACHIEVEMENT)			
Intercept	292.42	0.87	336.32**
Percent African-American	-7.45	0.87	-8.53**
Percent Hispanic	0.60	0.76	0.79
Average SES	9.01	1.13	7.98**
Students feel math is useful	-2.73	1.13	-2.41*
Students enjoy and feel competent in math	5.47	1.18	4.63**
Students disagree that math is more for boys	-0.46	0.95	-0.48
Percent of students on academic/college prep	4.71	1.09	4.31**
Mean years 12th graders have taken calculus	2.51	0.90	2.80**
GENDER BETA COEFFICIENT			
Intercept	-6.54	1.42	-4.59**
Percent African-American	1.52	1.75	0.87
Percent Hispanic	-0.81	1.23	-0.66
Average SES	1.26	2.74	0.46
Students feel math is useful	2.08	1.70	1.22
Students enjoy and feel competent in math	-2.89	1.80	-1.61
Students disagree that math is more for boys	-0.86	1.30	-0.66
Percent of students on academic/college prep	-0.58	1.77	-0.33
Mean years 12th graders have taken calculus	-0.98	1.57	-0.62
Mean composite math score of grade sample	2.05	3.58	0.57
RACE-ETHNICITY BETA COEFFICIENT			
Intercept	-13.76	2.11	-6.52**
Percent African-American	0.35	3.43	0.10
Percent Hispanic	1.02	1.85	0.55
Average SES	-4.03	3.12	-1.29
Students feel math is useful	-2.48	2.78	-0.89
Students enjoy and feel competent in math	-3.07	2.88	-1.07
Students disagree that math is more for boys	-3.44	2.06	-1.67
Percent of students on academic/college prep	-5.34	3.29	-1.62
Mean years 12th graders have taken calculus	-2.12	2.16	-0.98
Mean composite math score of grade sample	6.58	4.53	1.45
SES BETA COEFFICIENT			
Intercept	7.43	1.31	5.67**
YEARS OF GEOMETRY BETA COEFFICIENT			
Intercept	23.63	0.62	37.87**

Random within-school parameters	Reliability[4]	Parameter variance (Tau)[5]	Degrees of freedom	Chi-square test of Tau > 0[6]
INTERCEPT (AVG. ACHIEVEMENT)	0.69	65.98	128	413.58**
GENDER BETA COEFFICIENT	0.21	38.03	127	144.84†
RACE-ETHNICITY BETA COEFFICIENT	0.24	86.81	127	170.10*

[1]Average of five Gamma values. See technical notes for more information.
[2]Average of five standard error values plus standard error of the five Gammas. See technical notes for more information.
[3]Gamma divided by standard error. Probabilities based on a two-tailed test.
[4]Parameter variance divided by total variance. Average of five reliability values.
[5]Average of five parameter variance values.
[6]Average of five Chi-square tests.

NOTE: ** probability ≤ .01; * probability ≤ .05; † probability ≤ .10.

SOURCE: U. S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress, 1990 NAEP. Math Assessment of 12th Grade Students, Restricted-Use Data Base.

ISBN 0-16-045449-2
90000
9 780160 454493

United States
Department of Education
Washington, DC 20208-5653

Official Business
Penalty for Private Use, $300

Postage and Fees Paid
U.S. Department of Education
Permit No. G-17

Fourth Class Special
Special Handling

END

U.S. Dept. of Education

Office of Educational
Research and Improvement (OERI)

ERIC

Date Filmed
July 25, 1995

NOTICE

REPRODUCTION BASIS

☐ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☒ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").