DOCUMENT RESUME

ED 377 059 SE 055 439

AUTHOR	Mullis, Ina V. S.; And Others
TITLE	Effective Schools in Mathematics: Perspectives from the NAEP 1992 Assessment. Research and Development Report.
INSTITUTION	National Assessment of Educational Progress, Princeton, NJ.
SPONS AGENCY	National Center for Education Statistics (ED), Washington, DC.
REPORT NO	ISBN-0-16-045331-3; NAEP-23-RR-01; NCES-94-701
PUB DATE	Oct 94
NOTE	116p.
AVAILABLE FROM	U.S. Government Printing Office, Superintendent of Documents, Mail Stop: SSOP, Washington, DC 20402-9328.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	Elementary Secondary Education; *Family Influence; Grade 4; Grade 8; Grade 12; *Mathematics Achievement; *Mathematics Instruction; Questionnaires; *School Effectiveness; *Socioeconomic Influences; *Students
IDENTIFIERS	Mathematics Education Research; *National Assessment of Educational Progress

ABSTRACT

This research report provides information about the school context for learning mathematics and highlights some factors that National Assessment of Educational Progress (NAEP) analyses have found to be associated with effective schooling. Approximately 26,000 4th, 8th, and 12th graders in 1,500 public and private schools participated in the national assessment of mathematics. Data reported from questionnaire responses of students and school principals relate to school socioeconomic and demographic characteristics, student absenteeism, students changing schools, school problems and climate, high schools where students are college bound, impetuses for curriculum and instructional change, home support for academic achievement, mathematics classroom instruction, tracking, and course taking in grades 8 and 12. The most effective schools had students who watched less television, changed schools less often, were subject to only a moderate amount of testing in their mathematics classes (weekly to monthly), took more advanced courses, had positive attitudes toward academics, had fewer problems in the schools, and did mathematics and used calculators more frequently. The effectiveness of private schools was similar to that of public schools in which students, teachers, and parents have positive attitudes toward academics, and where few problems exist. At grades 4 and 8, a more stable student body with students who changed schools fewer times was associated with higher school effectiveness. In summary, students' home background and school socioeconomic indicators were powerful influences on academic achievement in mathematics. (MKR)

NATIONAL CENTER FOR EDUCATION STATISTICS

RESEARCH AND DEVELOPMENT REPORT

Effective Schools in Mathematics

Perspectives from the NAEP 1992 Assessment

THE NATION'S REPORT CARD naep

Prepared by Educational Testing Service under contract with the National Center for Education Statistics

Office of Educational Research and Improvement
U.S. Department of Education

What is The Nation's Report Card?

THE NATION'S REPORT CARD, the National Assessment of Educational Progress (NAEP), is the only nationally representative and continuing assessment of what America's students know and can do in various subject areas. Since 1969, assessments have been conducted periodically in reading, mathematics, science, writing, history/geography, and other fields. By making objective information on student performance available to policymakers at the national, state, and local levels, NAEP is an integral part of our nation's evaluation of the condition and progress of education. Only information related to academic achievement is collected under this program. NAEP guarantees the privacy of individual students and their families.

NAEP is a congressionally mandated project of the National Center for Education Statistics, the U.S. Department of Education. The Commissioner of Education Statistics is responsible, by law, for carrying out the NAEP project through competitive awards to qualified organizations. NAEP reports directly to the Commissioner, who is also responsible for providing continuing reviews, including validation studies and solicitation of public comment, on NAEP's conduct and usefulness.

In 1988, Congress created the National Assessment Governing Board (NAGB) to formulate policy guidelines for NAEP. The board is responsible for selecting the subject areas to be assessed, which may include adding to those specified by Congress; identifying appropriate achievement goals for each age and grade; developing assessment objectives; developing test specifications; designing the assessment methodology; developing guidelines and standards for data analysis and for reporting and disseminating results; developing standards and procedures for interstate, regional, and national comparisons; improving the form and use of the National Assessment; and ensuring that all items selected for use in the National Assessment are free from racial, cultural, gender, or regional bias.

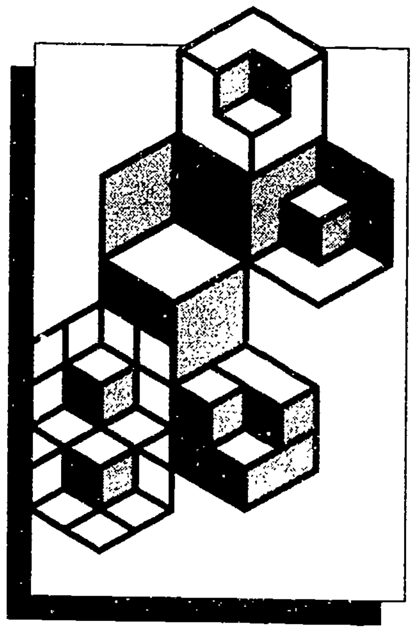
The National Assessment Governing Board

NATIONAL CENTER FOR EDUCATION STATISTICS

RESEARCH AND DEVELOPMENT REPORT

Effective Schools in Mathematics

Perspectives from the NAEP 1992 Assessment



Ina V. S. Mullis

Frank Jenkins

Eugene G. Johnson

Report No. 23-RR-01

October 1994

THE NATION'S REPORT CARD

Prepared by Educational Testing Service under contract with the National Center for Education Statistics

Office of Educational Research and Improvement
U.S. Department of Education

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

Education Assessment Division
Gary W. Phillips
Associate Commissioner

FOR MORE INFORMATION:

For ordering information on this report, write:

Education Information Branch
Office of Educational Research and Improvement
U.S. Department of Education
555 New Jersey Avenue, NW
Washington, D.C. 20208-5641

or call 1-800-424-1616 (in the Washington, D.C. metropolitan area call 202-219-1651).

Library of Congress, Catalog Card Number: 94-68491

ISBN: 0-88685-165-3

The work upon which this publication is based was performed for the National Center for Education Statistics, Office of Educational Research and Improvement, by Educational Testing Service.

CONTENTS

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 of studies that are, to some extent, on the "cutting-edge" of methodological developments. Emerging analytical approaches and new computer software development 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 three goals is that these reports present results or discussions 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 directed to:

Susan W. Ahmed
Acting Associate Commissioner
Statistical Standards and Methodology Division
National Center for Education Statistics
555 New Jersey Avenue NW
Washington, DC 20208-5654

Introduction

The purpose of this research report is to provide information about the school context for learning mathematics in our nation and to highlight some factors NAEP analyses have found to be associated with effective schooling. Because mathematics performance functions as an important gatekeeper in determining which students will go to college and join the professions,[1] it is of particular interest to examine educational contexts for this curriculum area.

As part of the 1992 mathematics assessment, NAEP collected questionnaire data from students as well as their teachers and school administrators. The *1992 NAEP Background Questionnaire Framework*[2] focused on five major educational policy areas — instructional content, instructional practices and experiences, teacher characteristics, school conditions and context, and conditions outside of school that affect learning and instruction.

[1] Pelavin, S. & Kane, M., *Changing the Odds: Factors Increasing Access to College* (New York, NY: College Board Publications, 1990).

[2] National Assessment of Educational Progress, *1992 Background Questionnaire Framework* (Princeton, NJ: National Assessment of Educational Progress, Educational Testing Service).

This report provides the national percentages for various response categories to some of the student and school NAEP background questions, with a particular focus on those questions pertaining to school conditions and context. Approximately 26,000 fourth, eighth, and twelfth graders in 1,500 public and private schools across the country participated in the national assessment of mathematics. Additionally, the results are presented for the top-performing one-third of schools (as judged by performance on the NAEP mathematics assessment) compared to the bottom-performing one-third of schools.

While comparisons between performance and practices in the top- versus bottom-performing one-third of the schools present one view of school effectiveness, another approach is to study the effectiveness of schooling as the contribution it makes to learning beyond students' home backgrounds and the socioeconomic levels of the communities in which their schools were located. Thus, for the student and school questionnaire data, NAEP did additional analyses using hierarchical linear modeling (HLM) techniques to adjust for associations among key schooling variables and socioeconomic status. Although such adjustments are, for methodological reasons, always incomplete, this additional perspective helps underscore the role of some of the home and school factors examined by NAEP.

The following sections of this introduction discuss some of the issues involved in collecting and reporting large-scale questionnaire data. The discussions also include brief summaries of both the response data and the results of the hierarchial linear analyses of these data.

The School Questionnaire Results

The selection of background questions included in the NAEP mathematics assessment was guided by the *NAEP 1992 Background Questionnaire Framework*. In developing the framework, the committee of policy analysts and researchers considered the wide body of available research about factors influencing student learning as well as the particular purposes and strengths of NAEP data collection. In contrast to other questionnaires which may ask for detailed information on a small set of topics, the 1992 NAEP questionnaires attempted to include a few well-targeted questions on a wide variety of topics. Thus, a limited pool of questions was available and the results necessarily must be considered in light of ongoing research about school practices.

From a research perspective, the inclination is towards in-depth examinations of discrete topics. However, NAEP always has been sensitive to its voluntary nature and the burden it places on its respondents — students, teachers, and principals — who have very busy lives and receive nothing in return for the time they give to NAEP. Thus, an effort has been made to curtail the questionnaires to about 15 to 30 minutes for any one respondent (students, of course, also spend approximately an hour on the achievement portion of the assessment). When there are constraints, such as the operational, methodological, and budgetary considerations faced by NAEP, choices need to be made. For NAEP in 1992, the choice was to collect some information about a broad range of topics highlighted in the research literature, rather than in-depth information about relatively few topics. In this way, the survey data collected by NAEP can be used to monitor trends in key areas, while ongoing research can be used to determine the school and home contexts that best foster academic achievement. NAEP is able to ask questions that collect information about whether schools and parents are implementing those strategies that research has found to be most effective in helping students learn.

It is unwise, however, to rely too heavily on responses to any single question. Also, particularly given the reform environment, there undoubtedly is some language ambiguity inherent in the questions. Different terms have different meaning to different people depending on differences in contexts and environments. Still, the data can be used to point practitioners in directions of examining their own programs in greater detail. Perhaps more than anything, the data from the NAEP questionnaires can provide educators with a basis to confirm or illuminate patterns observed at the local level. They provide a back-up litmus test for checking the reasonableness of local findings. They also provide a source of questions and approaches for examining local policy issues, conducting local studies, and creating local initiatives to change practice.

There were very large performance differences within grade level between the achievement of students in the highest-performing one-third of the schools compared to the lowest-performing one-third of the schools. These are not small or extreme segments of the school population. For example, one-third of the twelfth graders in the nation represents more than one million students. It is a matter of concern, then, to find that the twelfth graders in the lower-performing third of the schools had lower average mathematics achievement than the eighth graders in the higher-performing one-third of the schools. Because it is considerably more difficult to implement reform in schools with large numbers of disadvantaged students,[3] it is of further concern to find the extent to which economically disadvantaged and minority students are concentrated in the lower-performing schools. Very few students in the top-performing one-third of schools (0 to 2 percent) were attending schools in disadvantaged urban communities and these students showed very little racial/ethnic diversity (84 to 88 percent White students across the three grades assessed by NAEP). At grades 4 and 8, about half the students in the lower-performing one-third of the schools were in schools where the majority of the students were participants in the subsidized lunch program.

Students in the top-performing schools have a greater opportunity to learn by virtue of being in school more often and with less mobility. In the top-performing schools there was less absenteeism, class cutting, tardiness, and transiency. The continuity of instructional approach that can be adopted in these schools simply is not possible in the lower-performing schools. By grade 12, all but 1 percent of the students in bottom-performing schools were in schools where principals considered absenteeism at least a minor problem.

In general, students in the top-performing schools have considerably more support for academic achievement than those in lower-performing schools. Considerable research has found schools with positive school climates to be more effective that those without any central purpose or goals, and that such schools can work for even the most disadvantaged students.[4] For the twelfth graders in the lowest-performing third of schools, 22 percent were in schools receiving negative ratings for students' attitudes towards achievement and 26 percent in schools receiving negative ratings for parental support. Unfortunately, consistent with their low opinions of academic endeavors these twelfth graders reported reading very little, either in school or at home, for their schoolwork. Only 36 percent reported doing as much as 10 pages of daily reading for their schoolwork.

The differences in academic emphases between top- and bottom-performing schools were accentuated most clearly in students' reports about their mathematics coursework. Amount of coursework represents a key indicator of students' opportunity to learn, because it not only shows access to the curriculum but also because differences in instructional approaches in low-track classes further exacerbate the discrepancies in students' access to

---

[3] O'Day, J.A. & Smith, M.S. "Systemic Reform and Educational Opportunity." In S.H. Furhman, editor, *Designing Coherent Policy: Improving the System* (San Francisco, CA: Jossey-Bass, 1993).

[4] Hill, P.T., Foster, G.E., & Gendler, T., *High Schools with Character* (Santa Monica, CA: The Rand Corporation, 1990).

engaging classroom experiences.[5] Students in top-third schools were more than twice as likely to be enrolled in algebra by the eighth grade (27 versus 13 percent). Considering the sequential nature of the mathematics curriculum, and the practice of tracking the more able students into more advanced coursework, it also is not surprising that more students in top-third schools had progressed further in the mathematics "pipeline."[6] More than twice as many of the twelfth graders in top-third schools (43 versus 18 percent) reported having taken Algebra III or Pre-Calculus. Almost all twelfth-grade students (90 percent) in top-third schools reported having studied geometry. In contrast, more than one-third of the twelfth graders (36 percent) in bottom-performing schools reported never studying this "gatekeeper" course, which most colleges are requiring prior to entrance.

School Effectiveness in Light of Students' Home Background and School Socioeconomic Level

In addition to examining the characteristics of schools with the highest mathematics achievement, it also is informative to examine factors associated with schools that are most effective in maximizing learning beyond the students' home background and the socioeconomic levels of the communities in which the schools are located. From a value-added perspective, schools with a more difficult educational task can be just as effective as those with fewer hurdles to overcome and NAEP used hierarchical analysis methods to examine this issue. Unfortunately, the state of the art in educational measurement is not nearly sophisticated enough to completely untangle the relationships among the myriad inputs to students' learning and provide definitive answers. No matter how sophisticated the approach, the adjustments can never be complete and such is the case in attempts to examine the effects of socioeconomic status. There are a number of technical issues that arise in the application of hierarchical linear analyses to NAEP data. These issues are discussed in the Procedural Appendix. Also, please note that the teacher questionnaire data were not included in the analyses for technical reasons (see Procedural Appendix).[7]

---

[5] Oakes, J., *Multiplying Inequalities: The Effects of Race, Social Class, and Tracking on Opportunities to Learn Mathematics and Science* (Santa Monica, CA: The Rand Corporation, 1990).

[6] *Everybody Counts: A Report to the Nation on the Future of Mathematics Education*, L. Steen, editor, (Washington, DC: National Research Council, National Academy Press, 1989).

[7] For the teacher questionnaire data, please see Dossey, J.A., Mullis, I.V.S., & Gorman, S., *How School Mathematics Functions* (Washington, DC: National Center for Education Statistics, 1994).

Nevertheless, two different analyses were conducted after average mathematics achievement was adjusted for variations in students' home background and the socioeconomic levels of the communities in which their schools were located. In one of these analyses, achievement was predicted based on school socioeconomic indicators. Schools with higher performance than their counterparts with similar socioeconomic characteristics were considered to be the most effective, and schools that did not perform as well were considered to be the least effective. An analysis of differences between the 15 percent most- and least-effective schools according to the NAEP background data showed that in the most effective schools, especially at the upper grades:

- students watched less television
- students were tested a moderate amount in mathematics class
- students were taking advanced mathematics courses
- students and parents had more positive attitudes toward academic achievement
- there were fewer school problems
- students read more for schoolwork each day (either at school or at home)
- students worked mathematics problems from textbooks more frequently
- more students went on to four-year colleges or universities

In another analysis, variables and sets of variables were examined to determine the major predictors of school effectiveness, with the following results:

- Important classroom factors included moderate testing in mathematics class, doing problems from textbooks more frequently, and more frequent calculator use.
- Important school characteristics were a positive school climate and limited problems. A student body with stability, as opposed to transiency, also was important.
- The largest effects were related to having more students taking more advanced courses. The number of advanced mathematics courses taken was the most powerful predictor of students' mathematics performance after adjusting for variations in home background.

- School-level socioeconomic indicators, such as community type and percentage of students participating in the subsidized lunch program, also were powerful predictors and were correlated with classroom, school, and course-taking factors.

A Note on Interpretations

Associations between NAEP background factors and mathematics achievement must necessarily be interpreted cautiously given the correlational nature of NAEP data and the complexity of the context in which learning takes place. For example, associations may result because certain practices lead to higher levels of achievement, or because of differential instruction in which teachers tailor what they do based on their perceptions of students' abilities. That is, some instructional strategies may be more effective or appropriate for higher-achieving students, while others are more suitable for lower-achieving students. In particular, various remedial techniques are often associated with lower average proficiency — not because these strategies in any way cause lower achievement, but because the poorer-performing students need special assistance.

Also, as observed in Chapter One of this report, relationships between achievement and contextual variables are affected by socioeconomic factors, which make it easier for wealthy school districts to provide more continuity in their instruction than poorer districts, which also have to contend with a variety of social problems attendant with poverty, which in turn tend to depress student achievement.

Thus, in Chapter Two, NAEP made special efforts to examine the independent importance of variables that are themselves highly correlated to socioeconomic status. The results highlight the relationships between various factors and achievement even after accounting for differences in students' home background and school communities, indicating that these practices are associated with achievement under a variety of circumstances. However, again, because of NAEP's survey design, extreme care should be taken in going beyond these findings to ascribe cause and effect relationships. It is especially unwarranted to assume that, by process of elimination, variables beyond those found to be significant in these regression analyses are not equally or perhaps even more important than the ones that were significant. Common sense, experience, and other knowledge about our nation's educational system are key in interpreting NAEP data.

The NAEP assessment results are most useful when they are considered in light of other knowledge about the education system, such as trends in instructional reform, changes in the school age population, and societal demands and expectations. References are provided to assist the reader in finding additional related information about the topics covered.

1

Characteristics of High-Achieving and Low-Achieving Schools in Mathematics

The purpose of this chapter is twofold, offering both a look at some characteristics of the highest- and lowest-achieving schools in mathematics, and providing a summary of the student and school response data that emerged as significant in the more complex hierarchical analyses presented in Chapter Two. Presenting a straightforward comparison of the characteristics of high- and low-performing schools provides one perspective about the contexts for learning school mathematics in our nation.

To examine the relationship between level of school performance and level of student performance, NAEP sorted schools by their students' average performance on the 1992 mathematics assessment, identifying the top one-third and bottom one-third of the schools. In this introductory chapter, response data from NAEP's school and student questionnaires are presented for the nation as well as for the top- and bottom-performing one-third of the schools. In Chapter Two, the results are presented for two different analyses conducted after average mathematics achievement was adjusted for variations in students' home background and the socioeconomic levels of the communities in which their schools were located.

Table 1.1 presents the average mathematics proficiency for fourth-, eighth-, and twelfth-grade students (attending both public and private schools combined) as well as for the top-performing and bottom-performing one-third schools. By definition, the average mathematics proficiency was higher in the top-performing schools, but it is informative to compare the extent of the performance differences and to compare some characteristics of these schools.

Table 1.1

Average Mathematics Proficiency for the Nation, the Top-Performing One-Third Schools, and the Bottom-Performing One-Third Schools, Grades 4, 8, and 12

	GRADE 4	GRADE 8	GRADE 12
Nation	218 (0.7)	268 (0.9)	299 (0.9)
Top One-Third	237 (0.8)	289 (1.3)	316 (1.1)
Bottom One-Third	196 (1.2)	245 (0.9)	279 (1.0)

The standard errors of the estimated percentages and proficiencies appear in parentheses. It can be said with 95 percent confidence for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details).

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

The difference in performance between students in the top one-third and bottom one-third schools can be illustrated by cross-grade comparisons. For example, average proficiency for twelfth graders in the bottom-performing one-third schools was lower than it was for eighth graders in the top one-third schools. Similarly, eighth graders in bottom-performing one-third schools displayed average performance more like fourth graders in top one-third schools than like their own grade peers in the top one-third of the eighth-grade schools. In making such comparisons, however, it should be understood that fourth, eighth, and twelfth graders were each given grade-appropriate materials.

School Socioeconomic and Demographic Characteristics

Table 1.2 presents the percentages of students in various school settings generally related to the socioeconomic standing, including the type of community in which the school is located, the percentage of students within the school participating in the subsidized lunch or nutrition program, and whether the school is public or private.

Table 1.2

Percentages of Students by School Socioeconomic Factors, Grades 4, 8, and 12

	GRADE 4			GRADE 8			GRADE 12		
	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools
Size and Type of Community									
Advantaged Urban	12 (1.8)	27 (4.3)	1 (0.8)	10 (1.8)	20 (4.7)	0 (0.2)	12 (2.1)	28 (4.9)	0 (0.4)
Disadvantaged Urban	9 (1.4)	0 (0.0)	26 (3.9)	9 (1.3)	0 (0.0)	24 (3.9)	10 (1.4)	2 (1.3)	27 (4.1)
Extreme Rural	12 (2.2)	13 (3.7)	14 (5.3)	9 (2.6)	6 (3.7)	13 (5.0)	12 (1.6)	6 (2.4)	16 (2.9)
Other	66 (3.0)	60 (4.6)	58 (5.4)	72 (3.1)	74 (5.8)	63 (6.3)	66 (3.0)	64 (5.6)	57 (5.0)
Percentage of Students in School Receiving Subsidized School Lunch									
Less than 25 Percent	50 (3.0)	82 (4.3)	10 (2.8)	63 (3.3)	93 (3.1)	21 (4.7)	78 (2.7)	97 (2.0)	45 (5.4)
26 to 50 Percent	26 (2.8)	13 (4.3)	31 (5.2)	18 (2.7)	6 (1.8)	27 (4.8)	14 (2.7)	2 (1.8)	30 (4.9)
51 to 75 Percent	11 (1.8)	2 (2.0)	29 (4.9)	11 (2.2)	0 (0.3)	30 (6.1)	6 (1.2)	1 (0.6)	16 (3.5)
More than 75 Percent	12 (1.9)	3 (1.9)	31 (4.3)	8 (2.0)	1 (1.2)	22 (4.7)	2 (0.9)	0 (0.0)	8 (3.0)
Type of School									
Public	87 (1.0)	80 (2.3)	93 (2.3)	89 (0.9)	78 (3.1)	95 (1.4)	87 (1.2)	75 (4.5)	94 (2.5)
Catholic	8 (0.7)	14 (2.0)	3 (0.7)	6 (0.7)	12 (2.5)	2 (0.9)	8 (1.3)	16 (3.7)	4 (2.1)
Other Private	4 (0.6)	6 (1.5)	1 (0.7)	5 (0.7)	10 (2.1)	3 (1.0)	4 (1.0)	8 (2.2)	2 (1.4)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent confidence for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Size and Type of Community. Students were classified by the type of community in which their schools were located and by principals' reports of the percentages of students in their schools whose parents were classified into various occupational categories. The advantaged urban category represents about 10 percent of the students at each grade attending schools in suburban and urban communities where students' parents had professional or managerial jobs. The disadvantaged urban category represents the 10 percent of the students attending schools in suburban and urban locales that had high proportions of the parents on welfare or not regularly employed. The extreme rural category includes the approximately 10 percent of students attending schools in the most rural areas, where many of the parents were farmers or farm workers. The 70 percent of students who did not fall into one of these three "extreme" community categories were classified as attending schools in "other" types of communities.

For higher-performing schools, disproportionately higher percentages of students were in advantaged urban communities, while the opposite was observed for lower-performing schools. For example, 27 percent of the fourth graders in top-performing one-third schools attended schools in advantaged urban communities and none were in schools in disadvantaged urban communities. Thirteen percent were in extreme rural schools, while the majority (60 percent) were attending schools in other types of communities. Conversely, about one-quarter (26 percent) of the fourth graders in the bottom-performing one-third schools were in schools in disadvantaged urban areas and very few were in schools in advantaged urban areas (1 percent). Fourteen percent of the fourth graders in the lowest-performing schools were in extreme rural areas and 58 percent were from other types of communities.

Percentage of Students in Subsidized School Lunch Program. As would be expected based on numerous studies relating achievement to socioeconomic level, at all three grades, far more students in bottom-third schools than top-third schools received subsidized lunches. In bottom-third schools, 60 percent of the fourth graders, 52 percent of the eighth graders, and 24 percent of the twelfth graders were in schools where more than half the students received subsidized lunches. In top-third schools, only 5, 1, and 1 percent, respectively, of the students were in programs where a majority of the student body received lunch subsidies. Across the nation, schools with fewer than one-quarter of their students on subsidized lunch programs had higher average mathematics proficiencies than schools with at least three-quarters of their students enrolled in such programs.

Type of School. While from 87 to 89 percent of the nation's students attended public school compared to only 11 to 13 percent in private schools, it can be seen that there was disproportionate representation within the top- and bottom-performing schools. Across the three grades assessed, the top-performing schools were composed of 75 to 80 percent public school students, with the remainder being private school students. In comparison, nearly all the students in the bottom-performing one-third of the schools, 93 to 97 percent, were in public schools. Table 1.3 presents demographic data for the highest- and lowest-performing one-third of the schools.

Table 1.3

Percentages of Students by School Demographic Factors, Grades 4, 8, and 12

	GRADE 4			GRADE 8			GRADE 12		
	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools
Region									
Northeast	21 (0.9)	26 (4.2)	17 (2.7)	22 (0.8)	25 (4.9)	20 (2.8)	24 (0.6)	30 (3.7)	17 (2.9)
Southeast	24 (0.9)	12 (2.3)	44 (4.5)	25 (0.7)	12 (3.3)	41 (3.0)	24 (0.6)	12 (3.2)	39 (4.6)
Central	27 (0.5)	37 (3.7)	12 (3.0)	25 (0.6)	38 (4.8)	12 (2.0)	25 (0.6)	30 (4.1)	9 (2.7)
West	28 (0.7)	24 (3.5)	27 (3.9)	28 (0.7)	26 (5.2)	27 (3.0)	27 (0.9)	28 (3.9)	35 (3.9)
Race/Ethnicity									
White	70 (0.2)	88 (1.4)	40 (2.9)	70 (0.2)	86 (1.4)	43 (2.2)	71 (0.6)	84 (1.1)	45 (3.6)
Black	16 (0.1)	3 (0.7)	38 (2.3)	16 (0.1)	4 (0.9)	35 (1.6)	15 (0.4)	5 (0.8)	33 (2.6)
Hispanic	10 (0.2)	4 (0.6)	18 (1.2)	10 (0.2)	5 (0.7)	18 (0.9)	10 (0.5)	5 (0.6)	18 (3.0)
Asian Pacific/Islander	2 (0.2)	3 (0.7)	2 (0.3)	2 (0.2)	4 (0.8)	2 (0.4)	4 (0.2)	5 (0.5)	3 (0.8)
American Indian	2 (0.2)	1 (0.2)	2 (0.3)	1 (0.2)	1 (0.2)	2 (0.4)	1 (0.1)	0 (0.3)	1 (0.2)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent confidence for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Region. All else being equal, it would be expected that much like the national distributions, approximately one-fourth of the students in both the top- and bottom-performing schools would have been from each of the four regions of the country. Indeed, at grade 4, the top-performing schools were composed of about one-fourth students from the Northeast and West (26 and 24 percent, respectively). However, 37 percent of the fourth graders in higher-performing schools were from the Central region and only 12 percent were from the Southeast. In contrast, a substantial proportion of fourth graders in the bottom-performing one-third schools were in the Southeast (44 percent), 27 percent were in the West, 17 percent in the Northeast, and 12 percent in the Central region. This imbalance also was observed at grades 8 and 12. At grade 12, in particular, students in the lower-performing schools were disproportionately from the Southeast (39 percent) and the West (35 percent).

Race/Ethnicity. At grade 4, 88 percent of the students in the top-performing one-third of the schools were White students, with only a few percent in other racial/ethnic classifications. In bottom-performing one-third schools, the percentages of White and Black students were approximately equivalent with about two-fifths from each category, and nearly one-fifth were Hispanic students. Similar to the results for top-third schools, small percentages of Asian/Pacific Islander and American Indian students were found in bottom-third schools. Approximately the same data were obtained at grades 8 and 12. The top-performing one-third of the schools had little racial/ethnic diversity (84 to 86 percent White students, with about 5 percent each for Black, Hispanic, and Asian/Pacific Islander students), while racial/ethnic composition in the lowest-performing one-third of the schools was from 43 to 45 percent White students, 3[illegible] to 35 percent Black students, 18 percent Hispanic students, and 1 to 2 percent American Indian students.

The socioeconomic and demographic composition of the top one-third of the schools compared to the bottom-performing one-third of the schools raises concerns from several perspectives.[8] According to 1990 census data, the number of poor school-aged children grew by 6 percent since 1980, and

[8] Hamburg, D., "Children of Urban Poverty: Approaches to a Critical American Problem," reprinted from the Carnegie Corporation of New York 1992 Annual Report (New York, NY: Carnegie Corporation of New York, 1993).

the national poverty rate for school-aged children increased.[9] While the number of poor White children declined and the Black school-aged poverty population showed little change, the number of poor Hispanic and Asian/Pacific Islander children grew. The number of poor school-aged children became increasingly concentrated in the Western and Southwestern regions. Although the total numbers of school-aged children declined in the Southeast between 1980 and 1990, many Southern states had the highest poverty rates in the nation. Finally, school-aged poverty became more concentrated in our nation's largest cities.

Further, research suggests that particularly as they progress into the higher grades, lower-performing students receive less effective mathematics instruction.[10] They have less extensive and less demanding programs available to them, and fewer opportunities to take the critical gatekeeping courses that prepare them for further mathematics study, such as algebra and geometry in grades 7 through 9 or calculus in high school. Also, that economically disadvantaged and minority students are more concentrated in the lower-performing schools may constrain their future as part of systemic educational reform. Smith and O'Day underscore the particular difficulty of implementing complex and ambitious reforms in schools with large numbers of disadvantaged students, given the challenges and the resources that are present.[11]

Student Absenteeism

Many schools are confronted with high absentee rates and frequent turnovers of the student body. These problems, which may by linked to socio-economic and demographic characteristics of the students and their families, have the potential to disrupt students' learning of the curriculum.[12]

---

[9] *School Age Demographics: Recent Trends Pose New Educational Challenges* (Washington, DC: United States General Accounting Office, 1993).

[10] Oakes, J., *Multiplying Inequalities: The Effects of Race, Social Class, and Tracking on Opportunities to Learn Mathematics and Science* (Santa Monica, CA: The Rand Corporation, 1990).

[11] O'Day, J. A. & Smith, M. S., "Systemic Reform and Educational Opportunity." In S. H. Furhman, editor, *Designing Coherent Policy: Improving the System* (San Francisco, CA: Jossey-Bass, 1993).

[12] *Characteristics of At-Risk Students in NELS:88* (National Center for Education Statistics, U.S. Department of Education, 1992).

As Table 1.4 shows, school principals reported fewer problems with absenteeism in top-performing schools. By grade 12, all but 1 percent of the students in bottom-performing schools were in schools where absenteeism was considered at least a minor problem. Mirroring this relationship between attendance and performance, substantially more students in bottom-performing one-third schools than in top one-third schools were reported absent each day. Also, there was a dramatic rise in absentee rates from the lower to the higher grades. While just 10 percent of the fourth-grade students were in schools with an absentee rate of 6 percent or higher per day, nearly half (47 percent) of the twelfth graders were in schools reporting this rate.

Table 1.4

Percentages of Students by Principals' Reports on Absenteeism, Grades 4, 8, and 12

	GRADE 4			GRADE 8			GRADE 12		
	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools
Degree Student Absenteeism is a Problem									
Serious or Moderate	13 (1.8)	1 (1.4)	30 (4.5)	26 (3.0)	8 (4.6)	45 (4.4)	46 (3.2)	24 (5.7)	77 (6.7)
Minor	51 (2.7)	51 (5.0)	52 (4.2)	49 (3.5)	58 (6.8)	43 (5.6)	42 (2.9)	55 (6.4)	22 (5.1)
Not a Problem	36 (2.2)	48 (4.9)	19 (3.4)	25 (2.8)	35 (5.7)	12 (3.4)	11 (1.5)	20 (3.5)	1 (1.3)
Percentage of Students in School Absent on an Average Day									
0-2 Percent	33 (3.4)	46 (6.1)	29 (4.8)	19 (2.6)	34 (6.3)	12 (3.9)	10 (2.0)	12 (3.2)	8 (3.4)
3-5 Percent	57 (3.7)	49 (5.7)	51 (5.4)	52 (3.8)	54 (6.5)	48 (3.8)	42 (3.5)	59 (5.5)	24 (5.0)
6-10 Percent	10 (1.7)	4 (2.3)	19 (4.0)	25 (3.0)	12 (4.9)	29 (4.6)	41 (3.6)	27 (5.6)	51 (5.9)
More than 10 Percent	0 (0.3)	0 (0.0)	1 (0.9)	4 (1.4)	0 (0.0)	10 (4.5)	6 (1.3)	2 (1.5)	17 (4.0)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent confidence for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Students Changing Schools

Because the stability of the student body can affect the continuity of instruction,[13] students at grades 4 and 8 were asked about the number of times they had changed schools during the last two years due to a change in address. The results presented in Table 1.5 reveal that 61 percent of the fourth graders and 79 percent of the eighth graders reported no changes in schools. At grade 4, fewer students in bottom-performing one-third schools than in top one-third schools reported no changes (50 compared to 70 percent, respectively) and more reported one or two changes (32 compared to 23 percent) and three or more changes (18 compared to 6 percent). At grade 8, fewer students in bottom-third schools than in top-third schools reported no changes (74 compared to 84 percent) and more reported one to two changes (21 compared to 14 percent).

Table 1.5

Percentages of Students Reporting Having Changed Schools During the Past Two Years Because They Changed Where They Lived, Grades 4 and 8

	GRADE 4			GRADE 8		
	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools
No Changes	61 (0.8)	70 (1.3)	50 (1.7)	79 (0.6)	84 (1.0)	74 (0.9)
One to Two Changes	27 (0.6)	23 (1.1)	32 (1.3)	18 (0.5)	14 (1.0)	21 (0.8)
Three or More Changes	12 (0.5)	6 (0.7)	18 (1.1)	4 (0.3)	2 (0.3)	4 (0.4)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent confidence for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

[13] *Elementary School Children: Many Change Schools Frequently, Harming Their Education* (Washington, DC: United States General Accounting Office, 1994).

Characteristics of At-Risk Students in NELS:88 (Washington, DC: National Center for Education Statistics, U.S. Department of Education, 1992).

School Problems

NAEP asked the principals (or their designees) in the schools the degree to which each of the following was a problem: student tardiness, students' cutting of classes, physical conflicts among students, teacher absenteeism, racial or cultural conflicts, and student health. Response options were "serious or moderate," "minor," and "not a problem."

As shown in Table 1.6, tardiness and class cutting were not reported as prominent problems at grade 4, but by grade 12 only 7 to 18 percent of the students were in schools where these were not problems. At twelfth grade, the most commonly reported problem was student tardiness, with nearly two-thirds of the students (64 percent) in bottom one-third schools attending schools where administrators reported this as a moderate to serious problem. Forty-three percent of the high school seniors in the bottom one-third schools attended schools where cutting classes was reported as a moderate to serious problem.

School administrators reported that only one-third of the nations' fourth graders were in schools where physical conflicts among students were not a problem. The corresponding results for grades 8 and 12 were 21 and 30 percent, respectively. About 10 to 19 percent of the students across the three grades attended schools where administrators reported physical conflicts among students to be a moderate to serious problem. For schools in the lower one-third of the performance range, these percentages rose to approximately one-fifth to one-third of the students. This finding is consistent with the results of Metropolitan Life's survey of violence in public schools, which found school violence problems to be much greater in urban schools and schools with low-achieving students.[14]

Although teacher absenteeism was not reported as much of a problem at grade 4 (two-thirds of the students were in schools where administrators reported it was not a problem), by grades 8 and 12 the majority of students were in schools where this was at least a minor problem. One-fourth of the twelfth graders in bottom one-third schools were in situations where teacher absenteeism was considered a moderate to serious problem.

Similarly, problems with racial and cultural conflicts increased in middle and high schools. Three-fourths of the elementary school students attended schools where administrators reported that racial and cultural

[14] Louis Harris and Associates Inc., "Metropolitan Life Survey of the American Teacher 1993: Violence in America's Public Schools," December 1993.

conflicts were not a problem, compared to 54 percent of the eighth graders and 58 percent of the twelfth graders.

Research indicates that students from disadvantaged socioeconomic backgrounds tend to have more health problems than economically advantaged students.[15] Across all three grades, school administrators reported that 14 to 17 percent of the students in bottom one-third schools had serious or moderate health problems. The range was lower, from 3 to 7 percent, for students in top one-third schools. Nationally, the majority of students at all three grades were in schools where students' health was at least a minor problem. From 8 to 10 percent of the students were in schools where students' poor health was a moderate to serious problem.

[15] McKey, R. H. et al, *The Impact of Head Start on Children, Families, and Communities* (Washington, DC: U.S. Government Printing Office, 1985).

Table 1.6

Percentages of Students by Principals' Reports on the Degree of School Problems, Grades 4, 8, and 12

	GRADE 4			GRADE 8			GRADE 12		
	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools
Student Tardiness									
Serious or Moderate	11 (1.0)	1 (1.0)	26 (4.1)	21 (2.8)	14 (4.8)	38 (5.0)	51 (3.2)	41 (5.4)	64 (5.3)
Minor	55 (2.9)	56 (5.1)	50 (4.4)	56 (3.5)	53 (6.8)	52 (4.8)	42 (3.0)	50 (5.9)	30 (5.1)
Not a Problem	33 (3.0)	43 (5.1)	24 (3.9)	22 (2.9)	33 (6.2)	10 (3.6)	7 (1.7)	9 (2.4)	6 (2.9)
Students' Cutting Classes									
Serious or Moderate	0 (0.0)	0 (0.0)	0 (0.0)	4 (1.5)	0 (0.0)	12 (4.2)	25 (2.8)	14 (3.6)	43 (5.1)
Minor	3 (1.0)	0 (0.0)	7 (2.6)	33 (3.6)	18 (4.4)	46 (6.8)	57 (3.0)	60 (4.6)	50 (5.3)
Not a Problem	97 (1.0)	100 (0.0)	93 (2.6)	63 (3.7)	82 (4.4)	43 (6.9)	18 (2.8)	25 (4.2)	7 (2.8)
Physical Conflicts Among Students									
Serious or Moderate	10 (2.0)	4 (2.2)	17 (4.2)	19 (2.5)	5 (2.7)	33 (5.5)	12 (2.2)	5 (2.6)	19 (4.0)
Minor	57 (3.4)	43 (5.9)	59 (5.4)	60 (3.0)	63 (6.2)	54 (5.8)	58 (3.5)	56 (5.7)	68 (4.6)
Not a Problem	33 (2.9)	52 (5.4)	24 (4.7)	21 (2.2)	32 (5.6)	12 (3.5)	30 (3.3)	40 (5.9)	13 (3.4)
Teacher Absenteeism									
Serious or Moderate	4 (1.3)	1 (1.0)	10 (2.7)	6 (1.6)	0 (0.0)	14 (3.8)	13 (1.9)	5 (2.3)	24 (5.0)
Minor	30 (3.0)	29 (4.9)	32 (4.5)	46 (3.5)	45 (6.5)	51 (5.8)	46 (2.7)	44 (4.8)	50 (5.4)
Not a Problem	66 (3.1)	69 (4.6)	58 (4.5)	48 (3.1)	55 (6.5)	36 (5.7)	42 (2.9)	51 (5.0)	27 (5.4)
Racial or Cultural Conflicts									
Serious or Moderate	1 (0.7)	0 (0.0)	3 (1.7)	3 (1.4)	0 (0.0)	4 (2.5)	3 (1.1)	1 (1.2)	6 (2.9)
Minor	23 (2.9)	14 (5.1)	29 (4.4)	39 (2.4)	32 (6.4)	43 (6.5)	43 (3.2)	39 (6.1)	51 (6.1)
Not a Problem	75 (2.9)	86 (5.1)	68 (4.5)	58 (4.0)	68 (6.4)	53 (6.5)	54 (3.3)	60 (6.1)	43 (6.4)
Student Health									
Serious or Moderate	10 (1.7)	3 (1.9)	17 (3.7)	10 (2.0)	7 (3.2)	17 (4.3)	8 (2.0)	4 (2.6)	14 (4.2)
Minor	49 (2.7)	49 (5.5)	49 (5.2)	46 (3.7)	42 (6.3)	46 (5.7)	46 (3.5)	47 (6.6)	53 (5.6)
Not a Problem	41 (3.1)	49 (5.9)	35 (5.2)	44 (2.8)	51 (5.0)	38 (5.7)	46 (3.6)	49 (6.7)	33 (5.4)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent confidence for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

School Climate

Comparisons among schools, particularly in parent, student, and teacher motivation, are complex undertakings, and no study can demonstrate which kinds of schools are most effective for all students. Considerable research, however, has found schools with positive climates to be more effective than those without any central purpose or goals. Recent research by the Rand Corporation in New York City schools found that public and private schools alike can make key features available to all students, and that such schools can work for even the most disadvantaged.[16] In *High Schools with Character*, special purpose and Catholic schools were described as having certain features that distinguished them from zoned public schools, including a focus on student outcomes and a central curriculum as well as a strong commitment to parenting and accountability to the people who depended on their performance.

Undoubtedly, across the nation, a variety of circumstances can be found. Given current trends in school age demographics toward more poor school-aged children and higher concentrations of these children in urban areas,[17] some school administrators may feel that they are working in relative isolation to address larger societal problems, while also trying to provide quality education.[18] In another series of questions, school administrators were asked about the prevailing climate in their school toward academic achievement. Specifically, they were asked to characterize the following on a four-point scale from very positive to very negative: morale of teachers, students' attitudes toward academic achievement, teachers' attitudes toward academic achievement, parental support for students and teachers, regard for school property, and relations between students and teachers.

The data shown in Table 1.7 reveal that virtually all fourth graders were reportedly attending schools with at least a somewhat positive climate about each of these factors. The percentages of fourth graders attending schools that received very positive ratings from school administrators ranged from 37 percent of fourth graders attending schools where students were judged to have positive attitudes towards academic achievement to

[16] Hill, P. T., Foster, G. E., & Gendler, T., *High Schools with Character* (Santa Monica, CA: The Rand Corporation, 1990).

[17] *School Age Demographics: Recent Trends Pose New Educational Challenges* (Washington, DC: United States General Accounting Office, 1993).

[18] Kozol, J., *Savage Inequalities, Children in America's Schools* (New York, NY: Crown Publishers, Inc., 1991).

75 percent in schools where the relationships between students and teachers were judged as being very positive. Fewer than two-thirds of the fourth graders (62 percent) were in schools where administrators rated teacher morale as very positive and fewer than half (46 percent) were in schools where administrators rated parental support for students or teachers as very positive.

By grade 12, from 2 to 12 percent of the students were in schools with negative or very negative atmospheres regarding these factors. Only approximately one-fifth were in schools where students had very positive attitudes toward academic achievement, about one-third were in schools receiving high ratings from administrators for teacher morale, parental support, or regard for school property, and about half were in schools where the relations between students and teachers were very positive.

Especially at grades 8 and 12, the climate in the higher-performing schools was much more positive than in the lower-performing schools.

Table 1.7

Percentages of Students by Principals' Reports on the Climate at School, Grades 4, 8, and 12

	GRADE 4			GRADE 8			GRADE 12		
	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools
Morale of Teachers									
Very Positive	62 (3.3)	68 (5.6)	52 (4.7)	35 (3.0)	46 (7.1)	36 (3.9)	31 (2.6)	38 (4.4)	28 (4.5)
Somewhat Positive	34 (3.2)	30 (5.3)	40 (5.3)	55 (3.6)	48 (7.6)	53 (4.1)	58 (2.8)	56 (4.5)	59 (4.3)
Somewhat or Very Negative	4 (1.3)	2 (1.7)	8 (2.7)	9 (2.2)	6 (3.4)	11 (2.9)	11 (2.0)	6 (2.0)	13 (3.3)
Students' Attitudes Toward Academic Achievement									
Very Positive	37 (3.0)	44 (5.6)	27 (4.1)	24 (2.8)	41 (4.8)	8 (2.5)	21 (2.4)	37 (5.0)	9 (3.1)
Somewhat Positive	61 (3.2)	56 (5.6)	67 (4.7)	58 (3.2)	52 (6.3)	66 (4.9)	69 (2.7)	62 (5.0)	68 (4.8)
Somewhat or Very Negative	2 (0.9)	0 (0.0)	6 (2.3)	18 (2.6)	7 (3.0)	27 (4.5)	10 (1.7)	0 (0.3)	22 (4.4)
Parental Support for Students and Teachers									
Very Positive	46 (2.9)	57 (4.8)	28 (4.4)	33 (2.6)	56 (5.5)	14 (3.5)	33 (2.9)	53 (5.5)	14 (3.3)
Somewhat Positive	50 (2.9)	41 (4.8)	62 (4.2)	58 (2.8)	44 (5.5)	71 (5.9)	58 (3.1)	45 (5.8)	60 (6.1)
Somewhat or Very Negative	5 (1.2)	2 (1.4)	10 (2.7)	9 (1.7)	0 (0.0)	14 (4.3)	9 (1.8)	2 (1.6)	26 (5.1)
Regard for School Property									
Very Positive	49 (2.9)	59 (6.0)	36 (4.1)	34 (3.0)	41 (5.2)	19 (2.9)	32 (3.2)	44 (5.1)	21 (4.4)
Somewhat Positive	47 (2.9)	38 (6.0)	56 (4.3)	56 (2.9)	56 (5.2)	65 (5.3)	55 (2.8)	53 (5.3)	58 (4.4)
Somewhat or Very Negative	4 (1.1)	3 (2.5)	8 (2.4)	10 (1.9)	3 (1.8)	16 (3.5)	12 (2.0)	3 (1.8)	21 (4.0)
Relations Between Students and Teachers									
Very Positive	75 (1.8)	79 (3.3)	71 (3.0)	48 (3.2)	67 (1.6)	32 (4.0)	49 (3.8)	68 (5.9)	34 (5.7)
Somewhat Positive	25 (1.8)	21 (3.3)	29 (3.0)	51 (3.3)	33 (6.1)	66 (4.5)	49 (3.5)	32 (5.9)	63 (5.1)
Somewhat or Very Negative	0 (0.2)	0 (0.0)	1 (0.7)	1 (0.6)	0 (0.0)	2 (1.3)	2 (1.0)	0 (0.0)	3 (1.9)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent confidence for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

High Schools Where Students Are College Bound

As another indication of the academic orientation of each high school as well as another socioeconomic indicator, NAEP asked school administrators at grade 12 to provide the percentage of students from the previous graduating class who had gone on to attend a four-year college or university. The data are presented in Table 1.8.

For high schools across the nation, the higher the percentage of students who attended a four-year college after high school graduation, the higher the students' average mathematics proficiency. This finding is reinforced by the disparity in college attendance between students in top one-third and bottom one-third schools. While 60 percent of twelfth graders in top-third schools were in schools where the majority of the previous graduating class went on to attend a four-year college, just 13 percent of students in bottom-third schools were in a similar situation. For the majority of high school seniors in bottom one-third schools, one-fourth or fewer of the previous graduating class went on to attend a four-year college or university.

Table 1.8

Percentages of Students by Principals' Reports on the Percentage of Students from the Most Recent Graduating Class Who Attended A Four-Year College, Grade 12

Assessment Year-1992	25 Percent or Fewer	26 to 50 Percent	51 to 75 Percent	More than 75 Percent
Nation	26 (2.8)	36 (3.2)	26 (3.0)	12 (2.1)
Top One-Third	11 (3.4)	30 (5.7)	35 (4.8)	25 (4.8)
Bottom One-Third	55 (5.5)	32 (4.9)	10 (3.5)	3 (1.9)

The standard errors of the estimated percentages appear in parentheses. It can be said with about 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment.

Impetuses for Curriculum and Instructional Change

NAEP also asked school administrators about several possible impetuses to change curriculum and instructional practices within their schools. Given the three response categories of "to a great extent," "to some extent," and "not at all," school administrators described how much district or school testing programs, state testing mandates, public reporting of school or district performance data, and budget changes had served as an impetus to change during the past five years. As shown in Table 1.9, the general trend across the grades and questions was for approximately half of the students to be in schools where administrators responded that the various factors had served as an impetus to curriculum and instructional change only to some extent. The other half of the students were split fairly evenly between schools that were influenced to a great extent, and those schools that were not influenced at all during the past five years.

By grades 8 and 12, students in top one-third schools (31 to 43 percent) were about three times more likely than students in bottom one-third schools (8 to 13 percent) to be in situations where testing programs or mandates had not served as an impetus to curriculum or instructional change. However, these results do not necessarily indicate that testing programs and mandates cannot effectively guide reform; rather, they may show that schools with lower performance were more likely to see a need for testing programs than schools demonstrating higher performance.

As an impetus for change, public reporting of school or district performance data had similar results to testing programs or mandates as did the extent to which changes in the budget influenced curriculum and instructional revisions. Greater percentages of eighth and twelfth graders in top- than in bottom-performing schools were in schools where neither public reporting of performance data nor budget changes served as an impetus for curriculum and instructional change.

Table 1.9

Percentages of Students by Principals' Reports on Which Had Served as an Impetus to Change Curriculum and Instructional Practices Within the School, Grades 4, 8, and 12

	GRADE 4			GRADE 8			GRADE 12		
	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools
District or School Testing Program									
To a Great Extent	23 (2.8)	26 (5.1)	29 (4.4)	22 (3.0)	9 (3.5)	36 (6.8)	17 (2.5)	10 (3.2)	27 (4.7)
To Some Extent	58 (2.8)	57 (5.5)	51 (5.4)	52 (3.6)	49 (7.1)	53 (6.3)	58 (3.2)	59 (5.3)	61 (5.1)
Not at All	19 (2.0)	17 (3.3)	20 (4.1)	26 (3.2)	42 (7.4)	12 (3.7)	24 (3.0)	31 (4.0)	11 (3.6)
State Testing Mandates									
To a Great Extent	28 (3.0)	25 (4.5)	31 (5.3)	31 (3.5)	11 (4.9)	46 (6.5)	33 (2.9)	18 (5.0)	49 (4.9)
To Some Extent	50 (3.2)	51 (6.4)	53 (5.6)	45 (3.7)	46 (8.3)	42 (6.3)	42 (3.1)	38 (5.5)	43 (4.7)
Not at All	22 (2.2)	24 (4.8)	16 (4.1)	24 (3.0)	42 (8.8)	13 (3.4)	25 (3.1)	43 (6.3)	8 (2.3)
Public Reporting of School or District Performance Data									
To a Great Extent	19 (2.3)	19 (5.4)	23 (4.3)	19 (2.7)	5 (2.2)	26 (5.2)	17 (2.3)	11 (4.4)	28 (5.3)
To Some Extent	56 (3.5)	51 (7.6)	54 (6.3)	51 (4.1)	48 (7.1)	60 (5.6)	59 (3.2)	55 (5.7)	62 (5.8)
Not at All	25 (2.6)	30 (5.9)	23 (5.0)	30 (3.2)	47 (7.0)	14 (2.7)	23 (2.8)	34 (4.7)	10 (3.7)
Budget Changes									
To a Great Extent	19 (2.2)	19 (4.1)	25 (4.5)	19 (2.8)	13 (4.7)	23 (4.0)	31 (3.4)	25 (6.5)	39 (5.2)
To Some Extent	53 (3.2)	51 (5.4)	56 (5.1)	56 (3.3)	52 (6.8)	63 (5.2)	51 (3.4)	52 (7.1)	46 (5.6)
Not at All	28 (3.2)	30 (5.5)	19 (4.2)	25 (2.9)	35 (6.3)	14 (4.0)	18 (2.3)	23 (4.1)	15 (4.3)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent confidence for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Home Support for Academic Achievement

Since students only spend part of their day in school, Table 1.10 contains students' reports about several aspects of their home background. Not only are children's early sensibilities toward mathematics forged in the home, but parental support and encouragement can be instrumental in fostering higher achievement.[19]

[19] Stevenson, H. W. & Stigler, J. W., *The Learning Gap, Why Our Schools Are Failing and What We Can Learn from Japanese and Chinese Education* (New York, NY: Summit Books, 1992).

Comer, J. P., "Home, School, and Academic Learning." In J. T. Goodlad & P. Keating, *Access to Knowledge: An Agenda for Our Nation s Schools* (New York, NY: College Entrance Examination Board, 1990).

Table 1.10

Percentages of Students Reporting Various Aspects of Home Background, Grades 4, 8, and 12

	GRADE 4			GRADE 8			GRADE 12		
	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools
Number of Reading Materials in Home									
Zero to Two Types	30 (1.2)	16 (1.5)	44 (1.6)	20 (0.7)	10 (0.9)	32 (1.2)	15 (0.7)	10 (0.7)	24 (1.7)
Three Types	34 (0.7)	34 (1.1)	34 (1.2)	30 (0.7)	28 (1.5)	32 (1.0)	27 (0.8)	23 (1.4)	31 (1.4)
Four Types	36 (1.1)	50 (1.4)	22 (1.0)	50 (0.9)	63 (2.0)	35 (1.1)	58 (1.0)	67 (1.7)	45 (1.8)
Parents' Highest Level of Education									
Graduated College	41 (1.0)	54 (1.8)	31 (1.1)	42 (1.3)	58 (2.9)	27 (1.3)	43 (1.1)	59 (1.7)	29 (2.3)
Some Education After H.S.	7 (0.4)	6 (0.6)	7 (0.7)	18 (0.5)	17 (1.1)	17 (0.7)	26 (0.7)	23 (1.3)	27 (1.1)
Graduated High School	12 (0.5)	9 (0.8)	14 (0.9)	24 (0.7)	17 (2.1)	28 (1.9)	21 (0.8)	13 (1.1)	28 (1.5)
Did Not Finish H.S.	4 (0.3)	2 (0.5)	7 (0.5)	8 (0.5)	2 (0.6)	14 (1.2)	6 (0.4)	3 (0.5)	12 (1.5)
"I Don't Know"	35 (0.7)	28 (1.3)	41 (1.5)	9 (0.4)	6 (0.8)	14 (0.7)	3 (0.3)	2 (0.4)	4 (0.6)
Number of Parents in the Home									
Both Parents	77 (0.7)	85 (0.9)	66 (1.6)	76 (0.6)	84 (0.9)	66 (1.2)	77 (0.7)	82 (0.9)	66 (1.9)
Single Parent	19 (0.6)	12 (0.9)	27 (1.3)	21 (0.6)	15 (0.9)	29 (1.1)	19 (0.6)	15 (0.8)	26 (1.7)
Neither Parent	4 (0.3)	2 (0.4)	7 (0.8)	3 (0.2)	1 (0.3)	5 (0.5)	4 (0.3)	2 (0.4)	8 (0.7)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent confidence for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Reading Materials in the Home. Students were asked if magazines, daily newspapers, an encyclopedia, or 25 or more books were available in their homes Students in top-third schools were far more likely than those in bottom-third schools to report having all four types of reading materials available at home. For example, at grade 12, two-thirds of the students in top-third schools so reported compared to fewer than half (45 percent) in bottom-third schools.

Parents' Education Level. It should be noted that approximately one-third of the fourth graders did not know their parents' educational level. Still, across the three grades, students in top one-third schools (54 to 59 percent) were approximately twice as likely as students in bottom one-third schools (27 to 31 percent) to report that at least one parent had graduated from college. Conversely, while only 2 to 3 percent of students in top-third schools had parents who did not graduate from high school, 7 to 14 percent of students in bottom-third schools reported that their parents were not high school graduates.

Number of Parents in the Home. Approximately three-fourths of the students at each grade (76 to 77 percent) indicated that they had two parents residing at home. Substantially greater percentages of students in top one-third schools (82 to 85 percent) than in bottom one-third schools (66 percent at each grade) reported living at home with two parents, perhaps understandably, given the pressures associated with both child rearing and work.

In part, socioeconomic status can be linked to the number of parents living at home. A recent study by the U.S. General Accounting Office found that roughly half of all poor families were headed by a single parent.[20] The NAEP data closely correspond with 1990 census data showing that 25 percent of all families with children are headed by a single parent and that 59 percent of all Black children live with one parent.[21] It is predicted that these figures will continue to rise for the foreseeable future, particularly among families from disadvantaged socioeconomic backgrounds. The increase in single parenthood comes at the same time as research suggests that parents need to take a more active role in preparing their children for the rigors of school.[22]

Table 1.11 contains students' reports about how many pages they read each day in school and for homework across all of their subjects as well as their reports about how many hours of television they watched each day.

---

[20] *School Age Demographics: Recent Trends Pose New Educational Challenges* (Washington, DC: United States General Accounting Office, 1993).

[21] Maeroff, G. I., "Reform Comes Home: Policies to Encourage Parental Involvement in Children's Education." In C. E. Finn, Jr. & T. Rebarber, *Education Reform in the 90s* (New York, NY: Macmillan Publishing Company, 1992).

[22] "Parent and Preschool Factors That Influence Children's School Readiness," *ETS Policy Notes, 4(2)*, (Princeton, NJ: Educational Testing Service, 1992).

Table 1.11

Percentages of Students Reporting Number of Pages Read Each Day for Schoolwork and Hours of Television Watching Each Day, Grades, 4, 8, and 12

	GRADE 4			GRADE 8			GRADE 12		
	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools
Daily Reading for Schoolwork									
More than 10 Pages	56 (1.1)	63 (2.6)	49 (1.6)	40 (0.9)	45 (2.4)	37 (1.5)	46 (1.0)	56 (2.7)	36 (1.5)
Six to 10 Pages	22 (0.8)	20 (1.7)	23 (1.0)	29 (0.7)	31 (1.5)	26 (1.3)	24 (0.5)	21 (1.4)	26 (1.2)
Five or Fewer Pages	22 (0.9)	16 (1.5)	28 (1.5)	31 (0.9)	23 (2.0)	37 (1.7)	30 (0.9)	23 (1.7)	37 (1.6)
Daily Television Watching									
Six Hours or More	21 (0.7)	13 (1.3)	35 (1.4)	13 (0.4)	5 (0.5)	21 (1.1)	6 (0.4)	2 (0.4)	10 (0.8)
Four or Five Hours	22 (0.7)	22 (1.3)	20 (1.0)	25 (0.6)	18 (1.3)	31 (0.7)	15 (0.6)	10 (0.8)	23 (1.2)
Three Hours	17 (0.5)	18 (1.1)	15 (1.0)	22 (0.6)	22 (1.5)	20 (0.8)	19 (0.5)	16 (0.9)	21 (1.2)
Two Hours	19 (0.6)	22 (1.1)	13 (0.7)	24 (0.6)	31 (1.4)	16 (0.9)	27 (0.6)	29 (1.1)	23 (1.0)
One Hour or Less	21 (0.7)	26 (1.4)	17 (1.3)	17 (0.5)	23 (1.2)	12 (0.8)	33 (0.8)	43 (1.2)	23 (1.2)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent confidence for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Reading for Schoolwork. Considering a general perception that academic rigor increases as students progress through school, it is perhaps surprising that more fourth graders (56 percent) than eighth or twelfth graders (40 and 46 percent, respectively) reported reading more than 10 pages per day for their schoolwork. At both grades 8 and 12, 37 percent of students in bottom-third schools read five or fewer pages a day, compared to 23 percent of students in top-third schools.

Television Watching. Though many studies have linked excessive television viewing to lower scholastic achievement, the question of cause and effect remains. It may be that lower-achieving students simply tend to watch more television. Still, one study found that if students spent just half the time reading that they spent watching television, they could read over 100 additional books every year.[23]

Roughly 40 percent of the fourth and eighth graders reported watching at least four hours of television a day, although approximately half that many grade 12 students reported such frequent television viewing. Students in bottom-third schools were more likely than those in top-third schools to report this level of viewing.

Mathematics Classroom Instruction

In addition to the content and emphases placed on individual topics in the curriculum, the ways in which students are taught mathematics form the foundation for the inquiry, problem solving, and communication skills that students will need throughout their lives. The curriculum and its delivery can be viewed as the distribution system for the opportunity to learn mathematics.

As explained in the introduction, this report does not contain results from the teacher questionnaires administered as part of NAEP's 1992 mathematics assessment, because the small number of mathematics teachers per grade per school greatly complicates the multilevel analysis methodology used in Chapter Two. A companion report contains detailed information about mathematics instruction as described by students and their teachers.[24] However, Table 1.12 contains students' reports about several key variables, including solving problems from textbooks, using calculators, and frequency of testing. Also, please note that the grade 12 results are for the 64 percent of the students who reported being enrolled in a mathematics class.

[23] Kober, N., *EDTALK: What We Know About Mathematics Teaching and Learning* (Washington, DC: Council for Educational Development and Research, 1991).

[24] Dossey, J.A., Mullis, I.V.S., & Gorman, Steven, *How School Mathematics Functions* (Washington, DC: National Center for Education Statistics, 1994).

Table 1.12

Percentages of Students
Reporting Various Aspects of Mathematics Instruction, Grades, 4, 8, and 12

	GRADE 4			GRADE 8			GRADE 12		
	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools
Do Mathematics Problems from Textbooks									
Almost Every Day	66 (1.3)	67 (2.5)	66 (1.8)	85 (0.9)	91 (1.2)	79 (1.6)	88 (0.9)	88 (1.8)	86 (1.9)
At Least Once a Week	17 (0.9)	20 (1.8)	13 (1.3)	10 (0.7)	7 (1.0)	14 (1.3)	8 (0.6)	7 (1.1)	10 (1.3)
Less than Weekly	66 (1.8)	14 (1.1)	21 (1.4)	5 (0.4)	2 (0.5)	6 (0.7)	4 (0.6)	5 (1.1)	5 (1.0)
Use A Calculator									
At Least Weekly	21 (1.0)	23 (1.8)	22 (2.0)	53 (2.0)	62 (3.7)	39 (3.1)	82 (1.1)	86 (2.0)	73 (2.4)
Less than Once a Week	21 (1.2)	28 (2.7)	12 (0.8)	17 (0.8)	18 (1.4)	18 (1.3)	7 (0.6)	6 (0.9)	10 (1.3)
Never or Hardly Ever	58 (1.6)	48 (3.5)	66 (2.4)	30 (1.5)	20 (2.7)	43 (3.1)	11 (0.9)	8 (1.4)	17 (1.9)
Take Mathematics Tests									
Almost Every Day	9 (0.6)	5 (0.8)	15 (1.4)	6 (0.3)	3 (0.5)	10 (0.8)	4 (0.4)	3 (0.5)	6 (1.0)
At Least Once a Week	30 (1.2)	22 (1.9)	38 (1.8)	55 (1.2)	48 (2.2)	62 (1.8)	57 (1.4)	50 (2.1)	66 (2.6)
Less than Weekly	61 (1.5)	73 (2.1)	47 (2.3)	39 (1.3)	50 (2.5)	28 (2.0)	39 (1.5)	47 (2.1)	27 (3.1)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent confidence for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Doing Mathematics Problems from Textbooks. Approximately two-thirds of the fourth graders and most of the eighth and twelfth graders (85 to 88 percent) reported that they did problems from textbooks almost every day. At grade 8, greater percentages of students in top-third schools reported daily problem-solving from textbooks. At grade 4, fewer students in top-third schools reported doing problems from textbooks as an infrequent activity.

Using a Calculator. Parents and school administrators have tended to be cautious about, or even opposed to, implementing wider use of calculators in school classrooms. Some have expressed concern about damaging students' mastery of basic skills in mathematics.[25] In contrast, research shows that the proper use of calculators can enhance learning at all stages.[26] Calculators can take the drudgery out of mathematics and free the learner to concentrate on higher-order problem-solving skills. For example, *The NCTM Standards* make a clear statement supporting the important role calculators can play in helping students of all ages to explore, verify, and create mathematics.[27] *The NCTM Standards* call for all students to have access to appropriate calculators throughout their school experiences — in the lower grades a four-function calculator, a scientific calculator in the middle grades, and a graphing calculator thereafter. NAEP provided students with four-function calculators in grade 4 and scientific calculators in grades 8 and 12 for use in completing portions of the mathematics assessments.

Students' reports about the frequency of calculator use in mathematics class revealed increases across the grades assessed. About one-fifth of the fourth graders, one-half of the eighth graders, and four-fifths of the twelfth graders taking a mathematics class reported using calculators on at least a weekly basis. Between students in the top- and bottom-performing one-third schools, fourth graders reported little difference in calculator use. However, at grades 8 and 12, greater percentages of students in the top-performing schools reported using a calculator at least weekly.

Taking Mathematics Tests. While about two-fifths of the fourth graders reported taking mathematics tests at least weekly, about three-fifths of the eighth and twelfth graders reported being tested this frequently. At all three grades, students reported more testing in bottom one-third schools than in

---

[25] Campbell, P. F. & Stewart, E. L., "Calculators and Computers." In R. F. Jensen, *Research Ideas for the Classroom: Middle Grades Mathematics* (New York, NY: Macmillan, 1993).

Jensen, R. J. & Williams, B. S., "Technology: Implications for Middle Grades Mathematics." In D. T. Owens, *Research Ideas for the Classroom: Early Childhood Mathematics* (New York, NY: Macmillan, 1993).

[26] Lacampagne, C.B., *State of the Art, Transforming Ideas for Teaching and Learning Mathematics* (Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, Office of Research, U.S. Government Printing Office, 1993).

[27] National Council of Teachers of Mathematics, *Professional Standards for Teaching Mathematics* (Reston, VA: National Council of Teachers of Mathematics, 1991).

Kaput, J. J., "Technology and Mathematics Education." In D. A. Grouws, *Handbook of Research on Mathematics Teaching and Learning* (New York, NY: Macmillan, 1992).

top one-third schools. In both the 1990 and 1992 NAEP assessments, the frequency of testing was related to achievement, with more frequent testing generally associated with lower levels of student proficiency.[28] It may be that additional concern about the achievement of lower-performing students results in more testing, that additional time spent in testing detracts from ongoing classroom instruction, or some combination of these factors and others.

Tracking and Course Taking in Grades 8 and 12

Because students' primary opportunity to learn mathematics occurs during their schooling, there has been considerable concern about the amount and kinds of mathematics covered in the school curriculum, and about students' propensity to opt out of taking advanced mathematics coursework. Both the amount of time that students engage in learning and the quality of the instructional experience are potent and widely generalizable factors that have emerged from syntheses of educational research results.[29] This section summarizes NAEP data on eighth- and twelfth-grade course taking in mathematics.

Mathematics Coursework in Eighth Grade. Table 1.13 contains eighth graders' reports on the type of mathematics course they were taking at the time of the assessment as well as on which type of class they planned to take in ninth grade. Across the nation, about half the students were taking a general eighth-grade mathematics course, with another 28 percent enrolled in pre-algebra and 20 percent in algebra. Students in top one-third schools were more than twice as likely as those in bottom one-third schools to be enrolled in algebra by the eighth grade (27 versus 13 percent). Sixty-one percent of the eighth graders in bottom-third schools were in a general mathematics curriculum rather than pre-algebra or algebra, as opposed to 38 percent of the students in top-third schools.

---

[28] Dossey, J.A., Mullis, I.V.S., & Gorman, S., *How School Mathematics Functions* (Washington, DC: National Center for Education Statistics, 1994).

Mullis, I.V.S., Dossey, J.A., Owen, E.H., & Phillips, G.W., *The State of Mathematics Achievement: NAEP's 1990 Assessment of the Nation and the Trial State Assessment of the States* (Washington, DC: National Center for Education Statistics, 1991).

[29] Walberg, H.J., "Productive Time and Subject Matter Learning." In D.Tanner and J.W. Keefe, editors, *Improving the Curriculum: The Principal's Challenge* (Reston, VA: National Association of Secondary School Principals, 1988).

Table 1.13

Percentages of Eighth Graders Reporting on Mathematics Course Taking

	GRADE 8		
	Nation	Top One-Third Schools	Bottom One-Third Schools
What kind of mathematics class are you taking this year?			
Algebra	20 (1.0)	27 (2.4)	13 (1.4)
Pre-Algebra	28 (2.2)	32 (3.5)	22 (2.7)
Eighth-Grade Mathematics	49 (2.5)	38 (3.2)	61 (2.6)
Other Mathematics	3 (0.4)	3 (0.6)	4 (0.8)
What mathematics class do you expect to take in 9th grade?			
"I don't know"	21 (1.1)	15 (2.0)	28 (1.3)
Basic, General, Business or Consumer Mathematics	8 (0.7)	4 (0.7)	13 (1.1)
Pre-Algebra	15 (0.9)	12 (1.3)	18 (1.3)
Algebra I or Elementary Algebra	36 (1.4)	44 (2.3)	27 (2.0)
Geometry	14 (0.8)	19 (1.9)	11 (1.0)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent confidence for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

In part, the relationship between achievement and mathematics course taking is due to selection: The highest achieving seventh and eighth graders are identified as having the mathematical skills to study algebra I rather than taking a regular mathematics class. This tracking system then permits these students to move into geometry in grade 9. Significantly more eighth graders in top one-third schools than bottom one-third schools planned to take geometry (19 versus 11 percent) or algebra (44 versus 27 percent), with more students in bottom-third schools expecting to enroll in less advanced courses such as pre-algebra or general mathematics.

Because of the sequential nature of the mathematics curriculum and the practice of tracking, course taking has an enormous impact on students' opportunity to learn. Most obviously, coursework permits students' access to the curriculum. Those students who never take more challenging coursework are unlikely to learn advanced mathematical concepts in out-of-school settings. Beyond that, however, research indicates that instruction is qualitatively different in high- and low-track classes.[30] Students in low-track classes have less exposure to more challenging goals such as inquiry and problem-solving skills and less access to the teaching strategies that are most likely to generate interest and promote learning.

High Schools that Offer Advanced Mathematics Courses. Table 1.14 summarizes principals' reports about which advanced mathematics courses were taught in their schools. Nearly all twelfth graders (95 percent) were in schools that taught a course in pre-calculus, algebra III, or elementary functions, and most (86 percent) were in schools that offered trigonometry. In contrast, only one-third were in schools that offered a course in probability and statistics. Twice as many students in top one-third schools as in bottom one-third schools had access to a course in probability and statistics — 40 percent compared to 20 percent. Twenty-nine percent of the twelfth graders in top-third schools were in schools with one or more courses in AP calculus (26 or more students taking AP calculus) compared to 11 percent of the twelfth graders in bottom-third schools.

[30] Raudenbush, S.W., Rowan, B., and Cheong, Y.F., "Higher Order Instructional Goals In Secondary Schools: Class, Teacher, and School Influences," *American Educational Research Journal*, Vol. 30 (3), 1993.

Oakes, J., *Multiplying Inequalities: The Effects of Race, Social Class, and Tracking on Opportunities to Learn Mathematics and Science* (Santa Monica, CA: The Rand Corporation, 1990).

Table 1.14

Percentages of Students by Principals' Reports on Advanced Mathematics Courses Taught at Their Schools, Grade 12

YES, PRE-CALCULUS, ALGEBRA III, OR ELEMENTARY FUNCTIONS			YES, TRIGONOMETRY			YES, PROBABILITY/ STATISTICS		
Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools
81 (1.2)	90 (1.5)	62(3.1)	42 (1.5)	59 (2.2)	25 (2.2)	17 (0.7)	20 (1.5)	16 (1.2)

	None	1 to 5	6 to 10	11 to 25	26 to 50	More than 50
How many students in your school are currently enrolled in an Advanced Placement course in calculus?						
Nation	30 (3.1)	7 (1.5)	7 (1.7)	36 (3.0)	12 (2.1)	8 (1.5)
Top One-Third Schools	24 (4.7)	5 (2.9)	8 (2.8)	34 (5.3)	16 (3.4)	13 (3.4)
Bottom One-Third Schools	34 (6.4)	10 (3.1)	10 (3.6)	35 (5.8)	10 (3.7)	1 (0.6)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent confidence for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Mathematics Coursework in Twelfth Grade. It is informative to examine both when high school seniors first took algebra and how far they advanced in their algebra coursework. As shown in Table 1.15, 23 percent of the twelfth graders reported having taken algebra before the ninth grade, 51 percent reported taking algebra in the ninth grade, and most of the remainder took algebra in the tenth or eleventh grades. Six percent reported no study of algebra. Twelfth graders in top one-third schools were twice as likely as those in bottom one-third schools to have taken algebra before ninth grade (31 versus 15 percent).

Table 1.15

Percentages of Twelfth Graders Reporting on the Grade Level at Which They Initially Took a First-Year Algebra Course

	Before 9th Grade	9th Grade	10th Grade	11th or 12th Grade	Have Not Studied Algebra
Nation	23 (1.0)	51 (1.4)	15 (0.8)	5 (0.5)	6 (0.5)
Top One-Third Schools	31 (1.6)	54 (2.0)	11 (1.0)	2 (0.4)	2 (0.4)
Bottom One-Third Schools	15 (1.6)	48 (2.4)	19 (1.6)	8 (1.2)	9 (1.5)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent confider .or each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Table 1.16 contains the twelfth graders' reports on the extent of their schooling in algebra. Approximately two-thirds of the twelfth graders reported at least some study of Algebra II. Students were asked if they had studied Algebra II more than one year, one school year, one-half year or less, or if they had not studied it. "Some study" of the subject was defined as all but those students responding that they had not studied the course. Approximately four-fifths of the twelfth graders in top-third schools had studied Algebra II compared to only about half the students in bottom-third schools.

Only 30 percent of the twelfth graders nationwide reported advanced coursework, including some study of pre-calculus, third-year algebra, elementary functions, or analysis. Fourteen percent reported calculus coursework. Again, students in the higher-performing schools reported taking more advanced courses than their counterparts in lower-performing schools. As a point of comparison, 18 percent of the students in top-third schools reported calculus coursework, the same percent in bottom one-third schools that reported coursework in Algebra III.

Table 1.16

Percentages of Twelfth Graders Reporting at Least Some Study of Algebra II, Algebra III, and Calculus.

YES, ALGEBRA II			YES, ALGEBRA III OR PRE-CALCULUS			YES, CALCULUS		
Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools
68 (1.4)	82 (2.2)	54 (2.8)	30 (1.2)	43 (2.1)	18 (1.8)	14 (0.8)	18 (1.7)	10 (1.1)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent confidence for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

The role of geometry in the American educational system has changed over the years. Some educational researchers have cited geometry as the new "gatekeeper" course for access to higher education, since most colleges are now requiring the completion of a course in geometry prior to entrance.[31]

Table 1.17 shows the percentage of twelfth graders who reported at least some study of geometry as a separate course as well as the percentages who reported at least some study of trigonometry and statistics. Seventy-seven percent of the twelfth graders reported at least some study of geometry. Most students (87 percent) in top-third schools reported some geometry study compared to only 64 percent in bottom-third schools. Far fewer twelfth graders across the nation reported studying trigonometry as a separate course — 42 percent. Here, 59 percent of the twelfth graders in top-third schools reported some coursework compared to 25 percent in bottom-third schools. Although statistics was not a popular course, 17 percent of the twelfth graders reported having taken some statistics.

[31] Pelavin, S. & Kane, M., *Changing the Odds: Factors Increasing Access to College* (New York, NY: College Board Publications, 1990).

Table 1.17

Percentages of Twelfth Graders Reporting at Least Some Study of Geometry, Trigonometry, and Statistics.

YES, GEOMETRY			YES, TRIGONOMETRY			YES, STATISTICS		
Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools	Nation	Top One-Third Schools	Bottom One-Third Schools
81 (1.2)	90 (1.5)	62(3.1)	42 (1.5)	59 (2.2)	25 (2.2)	17 (0.7)	20 (1.5)	16 (1.2)

The standard errors of the estimated percentages appear in parentheses. It can be said with 95 percent confidence for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix for details). When the proportion of students is either 0 percent or 100 percent, the standard error is inestimable. However, percentages 99.5 percent and greater were rounded to 100 percent and percentages 0.5 percent or less were rounded to 0 percent. Percentages may not total 100 percent due to rounding error.

SOURCE: National Assessment of Educational Progress (NAEP), 1992 Mathematics Assessment

Despite the high goals set for mathematics achievement in the United States, the percentages of students taking advanced coursework remain low, yet, NAEP data reported elsewhere show that the trend is toward more course taking.[32] For example, between 1986 and 1992, trends in mathematics course taking for 17-year-olds, who are primarily juniors in high school, showed a significant decrease in the percentage of students who had only taken pre-algebra and a significant increase in those who reported having completed Algebra II.

[32] Dossey, J.A., Mullis, I.V.S., & Gorman, S., *How School Mathematics Functions* (Washington, DC: National Center for Education Statistics, 1994).

Mullis, I.V.S., Dossey, J.A., Campbell, Jay R., Gentile, Claudia, A., O'Sullivan, Christine, Latham, Andrew S., & Carlson, James, *NAEP 1992 Trends in Academic Progress* (Washington, DC: National Center for Education Statistics, 1994).

Summary

This chapter described the questionnaire response data for students and school principals (or their designees) for factors featured in the hierarchical linear analyses presented in Chapter Two. The data were presented both for the nation at grades 4, 8, and 12, and for the top-performing one-third of the schools compared to the bottom-performing one-third of the schools. (The top and bottom one-third schools were calculated by computing a mean for each school, then ranking and dividing the schools into thirds.) The results revealed a number of differences between the top-performing one-third of the schools and the bottom-performing one-third of the schools, both in their achievement and their characteristics.

For the same grade level, students in the top-performing one-third of the schools had considerably higher average mathematics proficiency than their counterparts in the bottom-performing one-third of the schools. For example, average proficiency for twelfth graders in the bottom-performing one-third schools was lower than it was for eighth graders in the top one-third schools. Further, economically disadvantaged students were concentrated in the lower-performing schools, which research shows will make implementing ambitious school reform in those schools quite challenging and difficult. Compared to the higher-performing schools, the lower-performing schools had more students from disadvantaged urban communities, more students participating in the subsidized school lunch program, and more racial/ethnic diversity. They also were more likely to be in the Southeast and to be public rather than private schools.

Absenteeism and student mobility were found to be higher in bottom-performing schools compared to top-performing one-third schools. Also, there was a dramatic rise in absentee rates from the lower to the higher grades. While just 10 percent of the fourth-grade students were in schools with an absentee rate of 6 percent or higher per day, nearly half (47 percent) of the twelfth graders were in schools reporting this rate.

Consistent with numerous other surveys and research, the principals in lower-performing schools reported having to deal with a number of social problems attendant with poverty. Especially at grade 12, compared to the counterparts in higher-performing schools, these principals reported more problems with student tardiness, cutting classes, physical conflicts among students, teacher absenteeism, racial or cultural conflicts, and student health. Tardiness and cutting classes, however, were reported as fairly widespread problems across all types of high schools.

Especially at grades 8 and 12, the climate in the higher-performing schools was much more positive than in the lower-performing schools. Again, however, by grade 12, principals reported that from 2 to 12 percent of the students were in schools with negative or very negative atmospheres regarding various school climate factors, including students' attitude toward academic achievement, teacher morale, parental support, regard for school property, and relations between students and teachers. Bottom-performing high schools reported fewer students who were college bound.

Across the grades, principals in the top-performing schools reported fewer impetuses to change curriculum and instructional practices within their schools, including the district or school testing program, state testing mandates, public reporting of school or district performance data, and budget changes. However, these results do not necessarily mean that testing programs and mandates cannot effectively guide reform; rather, they may show that schools with lower performance were more likely to see a need for testing programs than schools demonstrating higher performance.

The impact of the home was reflected in students' reports about several aspects of the home background. Compared to their counterparts in bottom one-third schools, students in top one-third schools reported more reading materials in the home, more highly educated parents, more likelihood of both parents living in the home, more daily reading for homework, and less television viewing.

Students in top-performing schools reported doing more mathematics problems from textbooks and more frequent use of the calculator as part of their mathematics instruction. They also reported less testing than students in the bottom-performing schools. Most importantly, however, they reported more advanced coursework than their counterparts in lower-third schools. Students in top one-third schools were more than twice as likely as those in bottom one-third schools to be enrolled in algebra by the eighth grade (27 versus 13 percent). Approximately four-fifths of the twelfth graders in top-third schools had studied Algebra II compared to only about half the students in bottom-third schools, and many more had taken geometry — 90 versus 62 percent — which is required by many colleges. Forty-three compared to 18 percent reported taking courses in Algebra III or pre-calculus. Only 14 percent of the twelfth graders nationwide reported studying calculus.

2

Examining Factors Associated with Effective Schools in Light of Students' Home Background and School Socioeconomic Level

Introduction

This chapter presents data based on using sophisticated hierarchical analytic techniques to further examine the relationships among the background factors described in Chapter One. In particular, given the degree of correspondence between a number of socioeconomic indicators and the top- versus the bottom-performing one-third of schools, it is important to emphasize that special efforts were made to study the impact of various school and instructional approaches after accounting for differences in students' home backgrounds and the socioeconomic levels of the communities in which their schools were located.

It is generally understood that learning results from many experiences, including those beyond schools and classrooms. Students' backgrounds and experiences outside of school also have a tremendous influence on final learning outcomes. Taken from this perspective, the effectiveness of schooling can be measured by the contribution it makes to learning beyond that obtained through external influences. That is, rather than simply looking at the characteristics of the schools with the highest mathematics performance, as was done in Chapter One, it is equally informative to examine factors associated with schools that appear to maximize learning beyond the students' home background and the socioeconomic levels of the communities in which the schools are located.

The NAEP data clearly illustrate that schools in the lower one-third of the mathematics proficiency distribution have many difficulties to overcome in comparison to schools in the top one-third of the proficiency distribution, including fewer students who are academically motivated, less stable student bodies resulting from more absences and families moving, and limited parental support for academic achievement. However, from a value-added perspective, schools with a more difficult educational task can be just as effective as those with fewer hurdles to overcome.[33]

To study the factors associated with effective schools, one common analytic approach has been to use regression techniques that try to make adjustments based on the varying levels of difficulty of the educational task. If schools were equal in terms of the difficulty of their educational task, what factors would make the difference? Which curricula, instructional practices, staff structures, and community supports would lead to the highest achievement?

Unfortunately, the technology associated with educational measurement is not nearly sophisticated enough to completely untangle the effects associated with the myriad inputs to student learning and provide definitive answers to these fundamental questions. Researchers working with large-scale databases that enable studying the effects of schooling in larger social contexts are hampered by the unwieldy aspects of those databases, and the interpretability of findings given the interrelatedness of the numerous factors involved and the hierarchies of decision making that affect educational practices.

[33] Cooley, W., "The Difficulty of the Educational Task," *The Pennsylvania Educational Policy Studies*, 1993.

Nevertheless, progress has continued in studying the factors associated with effective schools, especially since the 1966 release of *Equality of Educational Opportunity*, by James Coleman and others.[34] The findings of this controversial report indicated that differences among schools contributed little to differences in what students learned, and that, compared to home background, school had almost a negligible effect on differences in achievement. Since these results were contrary to common sense and considerable other evidence showing that students did learn from school, the report served as impetus to improving educational research methodology. Numerous subsequent studies have since confirmed that a variety of school factors can, in fact, make a difference.[35]

In some respects, however, problems plaguing the Coleman analysis and earlier researchers are still with us nearly 30 years later. Education is an incredibly complex undertaking, where home background prior to and during schooling has an enormous impact on schools and learning. Also, the effects are cumulative, with learning in early years having an impact on achievement in subsequent grades. Additionally, the naturally occurring hierarchies of the system, from students, classrooms, teachers, schools, districts, states, and up to the national level, all interact.

Foremost among the many difficulties in conducting school effectiveness research is the high degree of interrelatedness among all the processes involved in schooling. Because schools tend to serve communities, students attending the schools often have similar backgrounds. Similarly, the type of community will, to a certain degree, determine the resources available to the school. Within the schools, teachers and administrators will implement various approaches, but the effectiveness of these approaches (or lack thereof) could result from students' prior learning, their receptivity, teacher preparedness, having access to the necessary resources, having access to students to provide continuity of learning, and a host of other factors associated with schools in socioeconomically advantaged or disadvantaged communities.

---

[34] Coleman, J. S., et al., *Equality of Educational Opportunity* (Washington, DC: U.S. Department of Health, Education, and Welfare, U.S. Government Printing Office, 1966).

[35] For a broad sampling of such studies, see:

Wang, M. C., Haertel, G. D., Walberg, H. J., "Toward a Knowledge Base for School Learning," *Review of Educational Research*, Vol.63(3), 249-294, 1993.

What Works: Research About Teaching and Learning, Second Edition (Washington, DC: U.S. Department of Education, 1987).

Disentangling the effects of students' background and socioeconomic status from the effects of various different approaches to schooling becomes very complicated because of these interdependencies. In considering findings from studies that attempt to do so, it is impossible to ascribe cause and effect to single variables, in the sense that one single variable and no others will result in higher achievement. It also is extremely difficult to determine if the explanatory variables emerging from an analysis, no matter how powerful, are fundamental causes of phenomena or only related to other variables which are themselves fundamental causes. It would be unwarranted to expect that any analysis, including one encompassing a database as extensive as that of NAEP's 1992 mathematics assessment, would be able to specify the key to educational improvement. Longitudinal studies that follow the same students through their years of schooling have some capabilities for controlling for background factors and prior learning in determining the strength of school effects. However, even these studies are unable to completely untangle the web of causal priority involving individual, home, community, school, and classroom factors.

Yet, our country, as never before, has embraced the goals of raising academic achievement and providing the most effective education possible to all children.[36] Using comprehensive databases such as NAEP to examine various educational issues will contribute to the achievement of this goal. The variables under examination have been related to achievement in many previous studies, including longitudinal ones, and this previous research sets the groundwork for the current study. With each study, more information is added to our knowledge base about effective schools and more progress is made toward improving the methods for obtaining the most useful information from such databases. Using the NAEP database to examine the larger social contexts for education will not only help to illuminate those issues, but will provide an overview of the most significant relationships between various background factors and achievement, as a basis for further research.

[36] The National Council on Education Standards and Testing, *Raising Standards for American Education* (Washington, DC: U.S. Department of Education, 1992)

Hierarchical Analysis

One recent advance in educational research methodology has been the development of techniques that use multilevel linear models to analyze data nested within hierarchies.[37] For example, students may be found within classrooms, classrooms within schools, and schools within districts. By considering the variation of outcomes separately for each level of the hierarchy, the precision of estimates is improved. Also, by considering the nature of processes nested within groups, a greater richness of description is achieved. For example, in studying a school system, if student processes are held to be different for each school, we may estimate how school characteristics account for these differences. By looking at processes within and across hierarchies, we can gain understanding about how the educational system as a whole functions. (Information in the Procedural Appendix illustrates how single hierarchical linear models can be used to simultaneously estimate results across more than one hierarchical level.)[38]

Because the NAEP database contains a hierarchical structure — student, class, teacher, school, and state — analysts are beginning to apply hierarchical analytic methods to NAEP data to explore various issues.[39] NAEP data are particularly useful for examining effects across schools, because of the large number of schools that participate at each grade. Like other types of analyses based on regression, hierarchical linear modeling (HLM) can be implemented in various ways to answer various questions. The more extensive the database, as in the case of NAEP, the more numerous the possibilities for different applications.

---

[37] Raudenbush, S. and Bryk, A., "A Hierarchical Model for Studying School Effects," *Sociology of Education*, 59, 1-17, 1986.

Longford, N., "A Fast Scoring Algorithm for Maximum Likelihood Estimation in Unbalanced Mixed Models with Nested Random Effects," *Biometrika*, 74(4) 817-827, 1987.

Mason, W., Wong, G. & Entwisle, B., "Contextual Analysis Through the Multilevel Linear Model." In S. Leinhardt, *Sociological Methodology* 1983-1984, 72-103 (San Francisco: Jossey-Bass, 1984).

Goldstein, H. and McDonald, R., "A General Model for the Analysis of Multilevel Data," *Psychometrika*, 53(4), 455-467, 1988.

[38] Additionally, although not illustrated in this report, hierarchical analyses enable partitioning the variance among levels to indicate areas most fruitful for more detailed investigation. For an example of this type of three-level hierarchical analysis, please see: Raudenbush, S. W., Rowan, B., & Cheong, Y. F., "Higher Order Instructional Goals in Secondary Schools: Class, Teacher, and School Influences," *American Educational Research Journal*, Volume 30, Number 3, Fall 1993.

[39] Arnold, C., "Using Hierarchical Linear Models Using NAEP Data." Paper presented at the 1992 annual meeting of the American Educational Research Association, Atlanta, GA.

The results from two different applications of HLM analysis are presented in this chapter. Both of them were designed to provide further information about the effects of school factors beyond those associated with the socioeconomic level of the students' home background.[40] Both were conducted on the 1992 national mathematics assessment data at grades 4, 8, and 12, using data from two levels of organization — students within schools and schools themselves.[41]

To provide stability, the hierarchical analyses were based on schools having 15 or more sampled students. At grade 4, the sample consisted of 5,081 students in 224 schools. The student sample per school ranged from 15 to 55, with an average of 23 per school. At grade 8, the sample consisted of 4,979 students in 186 schools. The average number of students sampled per school was 27, ranging from 15 to 64. The twelfth grade sample consisted of 4,905 students in 189 schools. The average number of students sampled per school was 26, ranging from 15 to 52. (Further details are provided in the Procedural Appendix.)

Identifying Factors that Differentiate Between the Most Effective and Least Effective Schools

In the following application of HLM analysis, the most and least effective schools were identified and then compared with respect to the background variables collected by NAEP. Schools selected for membership in either the most or least effective group of schools were defined as those schools with average mathematics performance much higher or much lower than would be expected based on the home background of the students within the school and the overall socioeconomic level of the school (see Procedural Appendix).

The two levels in the multilevel model used for this analysis included a student level and a school level. At the student level, there were regressions based on data for the individual students within each of the schools to

[40] Jenkins, F., *Using Hierarchical Analyses to Identify Factors Contributing to Effective Schools*, Research Report, Educational Testing Service (in progress).

[41] Because there are typically few teachers of mathematics at each grade within a school, the teacher questionnaire data were not used in this analysis. Also, attempts to use three levels of organization, with states as the third level, were generally less successful because of the enormous size of the NAEP state-by-state database at a given grade — approximately 4,000 schools and 100,000 students. Full details of these HLM analyses will be covered in the forthcoming ETS Research Report by Jenkins, *Using Hierarchical Analyses to Identify Factors Contributing to Effective Schools*.

estimate mean school performance, controlling for students' home background. At the school level, regressions were conducted to predict schools' performance based on school socioeconomic level.

At the student level, *students' home background* was defined as a composite variable, based on three components: 1) students' reports about their access to reading materials in the home, 2) their parents' highest level of education, and 3) the number of parents living at home.

- *Reading materials in the home* was based on the sum of students' reports about receiving a newspaper regularly, an encyclopedia in the home, more than 25 books in the home, and receiving regular magazines.
- *Parents' highest level of education* was based on students' reports about the highest education level of either parent — graduated from college; some education beyond high school; and graduated from high school or less education (including students who did not know their parents' educational level).
- *Parents living in the home* was based on students' reports about whether both parents or stepparents, a single parent, or neither parent lived at home.

At the school level, *school socioeconomic level* also was defined as a composite variable based on three components: 1) the percentages of students receiving the subsidized school lunch and/or nutrition program, 2) the size and type of community in which the school was located, and 3) the home background of the students attending the school.

- The categories for percentages of students participating in the *subsidized school lunch* program were: 0 to 5, 6 to 25, and 26 to 100 percent.
- The *schools' size and type of community* was somewhat different by grade, with highest to lowest achievement associated with three categories at grade 4: 1) schools in small places, medium-sized cities, and advantaged urban areas; 2) extreme rural areas, rest of main big cities, and rest of urban fringes; and 3) disadvantaged urban areas. At grade 8, there were four categories: 1) advantaged urban; 2) small places; 3) extreme rural, rest of main big cities, rest of urban fringes, and medium cities; and 4) disadvantaged urban areas. At grade 12, the four categories were: 1) advantaged urban; 2) rest of main big

cities, rest of urban fringes, medium cities, and small places; 3) extreme rural; and 4) disadvantaged urban.

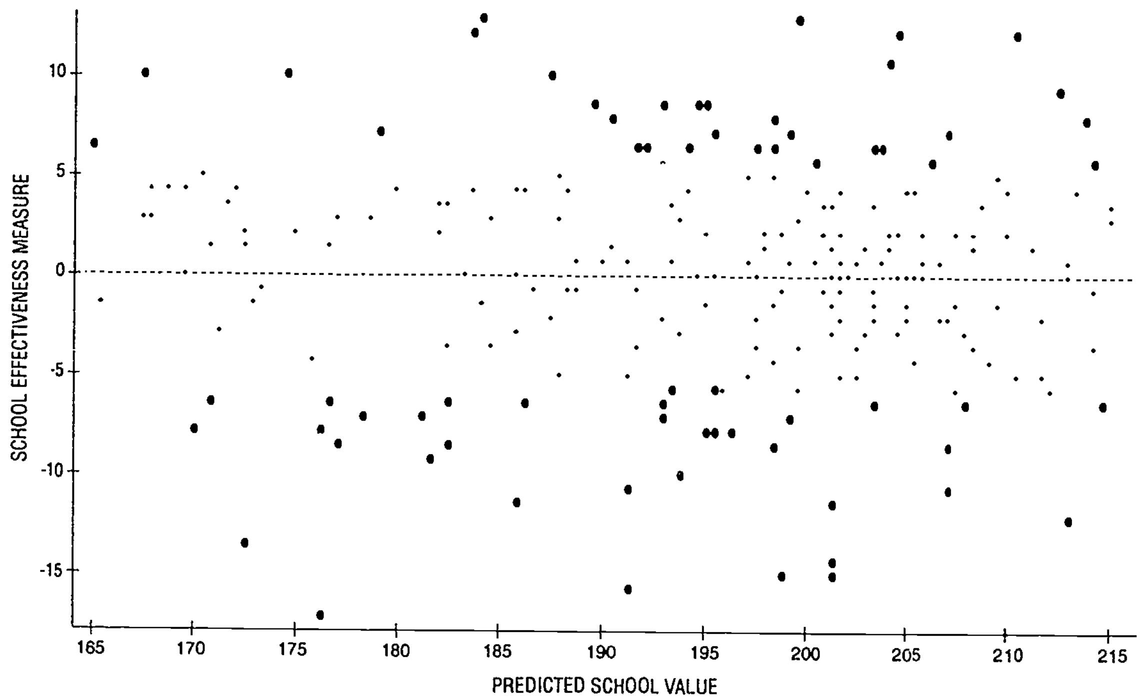
- *Students' home background* was defined as noted above at the student level, but aggregated to the school level.

Based on the HLM analyses, approximately 15 percent of the schools whose performances were considerably higher than predicted based on their socioeconomic level were designated as the most effective schools (i.e., schools at least one standard deviation above the average for schools at their socioeconomic level). Conversely, the approximately 15 percent of the schools with performances considerably lower than predicted were designated as the least effective schools (i.e., schools at least one standard deviation below the average for schools at their socioeconomic level). The two groups of schools — most and least effective — were then compared for any significant differences with respect to a large set of background factors (see Procedural Appendix).

Figure 2.1 plots the measure of school effectiveness against predicted mean school proficiency for each of the three grades assessed. The measure of school effectiveness was defined as the difference between the school mean adjusted for students' home background and the predicted school mean based on school socioeconomic characteristics. It can be seen that there are especially effective and ineffective schools across the complete range of predicted performance.

Figure 2.1

Plots of the School Effectiveness Measure Versus Predicted School Means by Grade

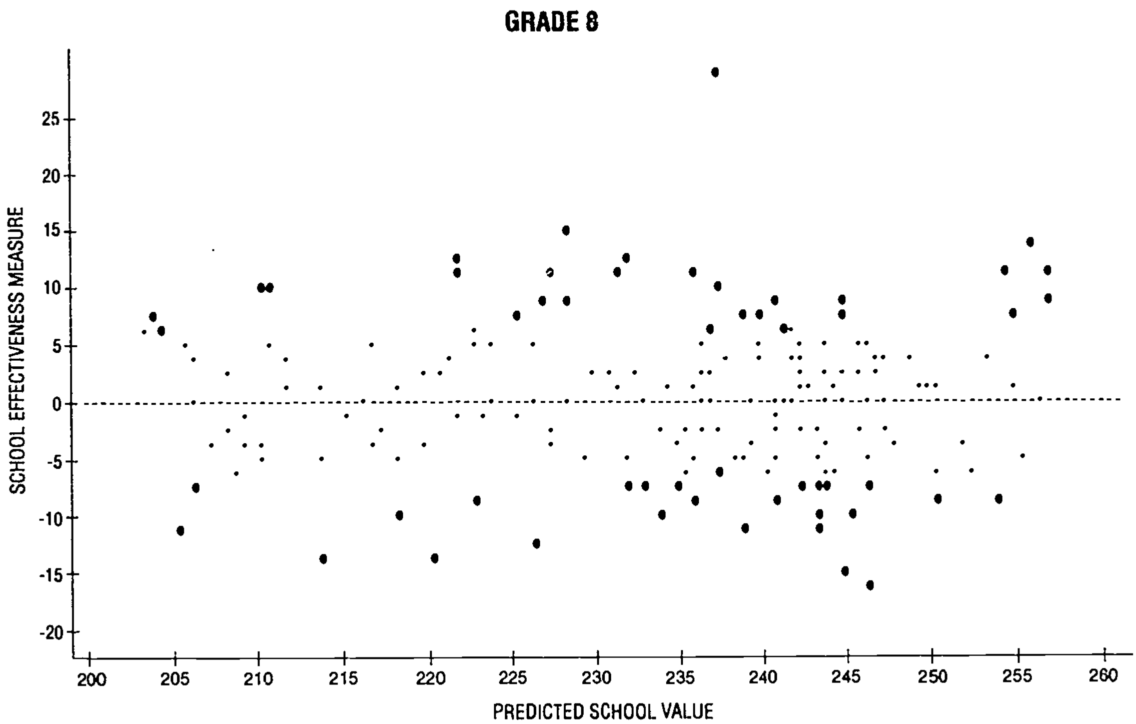


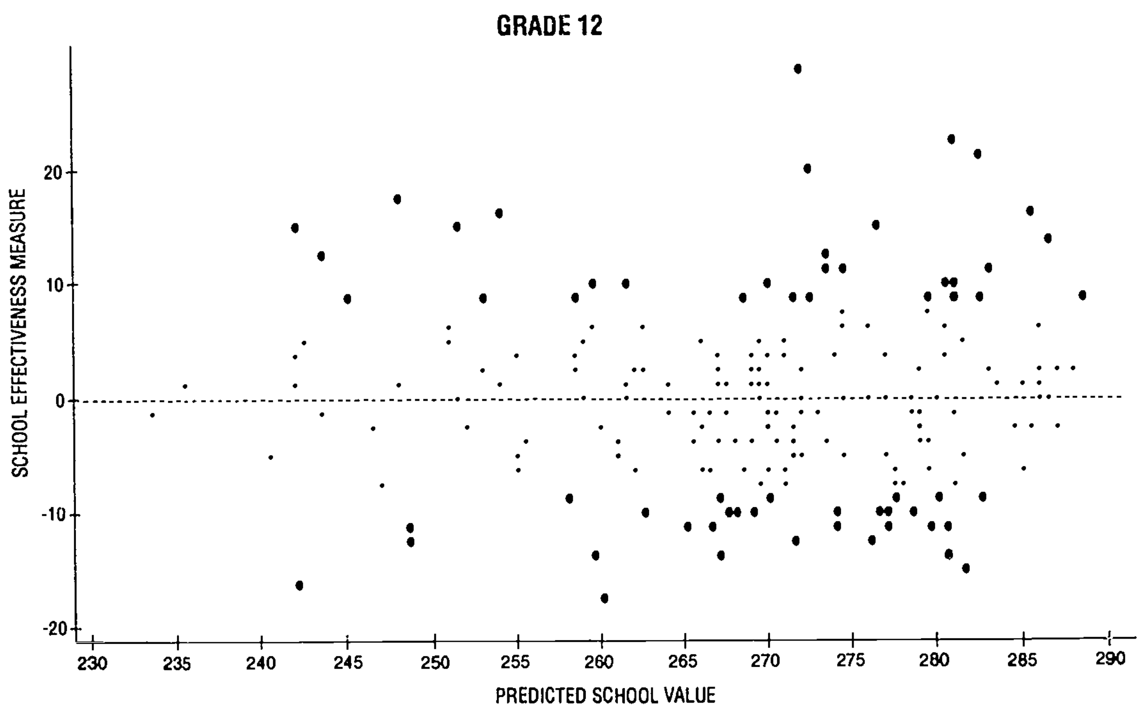
Note: Larger dots indicate schools one or more standard deviations from expected performance.

61

Figure 2.1 continued

Plots of the School Effectiveness Measure Versus Predicted School Means by Grade





Note: Larger dots indicate schools one or more standard deviations from expected performance.

Essentially, the extensive array of NAEP background variables was scrutinized in relation to mathematics proficiency using both correlational techniques and categorical approaches. Those variables with the most pronounced relationships were tested for significant differences between the most and least effective schools (see Procedural Appendix for details). Because this procedure included all factors related to proficiency, the analysis encompassed both instructional and demographic characteristics.

Variables with significant differences between the most and least effective schools at grades 4, 8, and 12 are presented in Table 2.1. At grade 4, six variables emerged as differentiating the most effective from the least effective schools.

Fourth graders in the most effective schools watched less television, as classified into three categories including the percentages watching 0 to 2 hours per night, 3 to 5 hours per night, and 6 or more hours per night. More students in the most effective schools were tested either once or twice a week or once or twice a month in mathematics class, rather than almost every day or hardly ever or never. Smaller percentages of the effective schools were in the Southeast versus in the Northeast, Central, or West. The most effective elementary schools had more stable student bodies, with smaller percentages of students having changed schools during the past two years (categorized as 0, 1, 2, and 3 or more changes). At grade 4, a larger percentage of public schools than private schools performed considerably higher than was typical for schools with similar socioeconomic characteristics. Finally, relations between students and teachers were more positive, as rated by school administrators on a four-point scale from very positive to very negative.

Table 2.1

Significant Factors Differentiating the Group of Most Effective Schools from the Group of Least Effective Schools

GRADE 4

Students watched less television.
Students changed schools fewer times in the past two years.
Students were tested weekly or monthly in mathematics class.
Smaller percentage of schools in the Southeast.
Relations between students and teachers were more positive.
Larger percentage of public schools.

These are the six variables showing significant differences between the most effective and least effective schools (at the .05 level). The multivariate test was significant at less than .0005. The canonical correlation was .64 (a correlation of 1.0 indicates that the values of these six variables would identify most and least effective schools with certainty).

GRADE 8

More students were currently enrolled in Algebra I.
More students planned to take geometry in ninth grade.
Students watched less television.
Students changed schools fewer times in past two years.
Students used calculators more frequently.
Smaller percentages of schools in the Southeast.
Larger percentages of White, Asian/Pacific Islander, and Hispanic students.
More schools where physical conflicts among students were not a problem.
Students had more positive attitudes toward academic achievement.
Parents provided more positive support toward student achievement.

These are the 11 variables showing significant differences between the most effective and least effective schools (at the .05 level). The multivariate test was significant at less than .003. The canonical correlations was .65.

GRADE 12

Students have taken more advanced mathematics courses.
Students took algebra before or in ninth grade.
More students read six more pages per day in school and for homework.
Students watched less television.
Students used calculators more frequently.
Students did mathematics problems from textbooks more often.
More students in last year's graduating class attending a four-year college or university.
Relations between students and teachers were more positive.
More schools where physical conflicts among students were not a problem.
Students had more positive attitudes toward academic achievement.
Parents provided more positive support toward student achievement.
More schools where student absenteeism was not a problem.
Teachers' morale was more positive.
Regard for school property was more positive.
Less use of testing data, state mandates, or budget information as impetus in making curricular or instructional change.

These are the 15 variables showing significant differences between the most effective and least effective schools (at the .05 level). The multivariate test was significant at less than .001. The canonical correlation was .73

At grade 8, variables related to having more students taking advanced mathematics courses started to differentiate between the most and least effective schools. The most effective schools had more eighth graders who planned to take geometry (or Algebra I) in the ninth grade, and more eighth graders enrolled in Algebra I (or pre-algebra). Both of these variables indicate a greater percentage of students in the higher tracks taking more advanced mathematics courses. Eighth graders taking Algebra I can take geometry in grade 9, while those in pre-algebra in eighth grade can take Algebra I in ninth grade. Perhaps related to these course-taking patterns, students in more effective schools were asked to use calculators more frequently in mathematics class (categories were: almost every day, once or twice a week, once or twice a month, never or hardly ever).

Similar to grade 4, eighth graders in the most effective schools watched less television, changed schools fewer times during the past two years, and were tested a moderate amount in mathematics class (once or twice monthly). Also, fewer of the more effective schools were in the Southeast. In contrast to grade 4, however, for the higher grades, being either a public or a private school (Catholic or other type of private school) was not significantly related to effectiveness.

The most effective schools at grade 8 were those where physical conflicts among students were less of a problem (on a scale: not a problem, minor, moderate, or serious), and there were larger percentages of White, Asian/Pacific Islander, and Hispanic students. Also, according to judgments of school administrators, the students were characterized as having more positive attitudes toward academic achievement, and parents provided more positive support toward student achievement (both of the latter variables were classified: very positive, somewhat positive, somewhat negative, and very negative).

At grade 12, a number of factors differentiated the most and least effective schools. In particular, an overall orientation toward academic achievement, sometimes referred to as the "press for academic learning,"[42] emerged at the high school level as a major factor in increased school effectiveness. That schools with a focused sense of purpose can be effective regardless of the economic levels of their students and communities is

[42] Oakes, J. "Tracking in Secondary Schools: A Contextual Perspective," *Educational Psychology*, 22, Spring, 1987.

consistent with school effectiveness research, including the Rand Corporation study of New York City high schools reported in *High Schools with Character*.[43]

After adjusting for the students' home backgrounds and the general socioeconomic level of the schools, the strongest differences evidenced by students in more effective high schools were:

- taking more advanced mathematics courses (geometry, Algebra II, trigonometry, pre-calculus/Algebra III/elementary functions, or calculus)
- reading at least six or more pages per day for school and homework
- doing mathematics problems from textbooks more frequently (almost every day, compared to once or twice a week, once or twice a month, or never/hardly ever)
- watching less television per day (categorized as 0 to 2 hours, 3 to 5 hours, and 6 hours or more)
- having very positive attitudes toward academic achievement (compared to somewhat positive, somewhat negative, or very negative attitudes)
- taking Algebra I either before or during the ninth grade

In addition, more school administrators in the most effective high schools reported that absenteeism was not a problem, nor were physical conflicts among students. Also, in the most effective high schools, higher percentages of students from last year's graduating class were now attending a four-year college or university. (For the last question, the six response categories were 0-10, 11-25, 26-50, 51-75, 76-90, and 91-100 percent). In general, the climate in these schools, as reported by school administrators, was more positive than in the least effective schools. In particular, administrators characterized the effective schools as having more positive climates with respect to higher teacher morale, more support from parents toward student achievement, better relations between students and teachers, and higher regard for school property.

[43] Hill, P. T., Foster, G. E., & Gendler, T., *High Schools with Character* (Santa Monica, CA: The Rand Corporation, 1990).

Finally, administrators reported that among the following four sources, fewer had served as an impetus to change curriculum or instructional practices: district/school testing programs, state mandates, public reporting of school/district performance data, and budget data. This latter finding may reflect less pressure or need for change in the more effective schools, or perhaps other sources of information were used to make decisions about change.

Examining the Relationships Among Various Factors Associated with Effective Schools

The purpose of the second — and more complicated — application of HLM analysis was to explore the relationship among factors associated with effective schools. The same two organizational levels were involved as in the analysis previously described — students within schools and schools themselves. Also, just as in that analysis, the purpose of the student level was to adjust schools' average mathematics proficiency for students' home background, and the same student level composite variable was used (reading materials in the home, parents' educational level, and number of parents in the home). The amount of student proficiency variance accounted for in the within-school adjustment for students' home background was modest: 11 percent for grade 4, 17 percent for grade 8, and 13 percent for grade 12. In this analysis, however, it is this adjusted school mean that was considered the measure of school effectiveness. At the school level of the HLM analysis, multiple regression techniques were used to enter single variables or sets of variables sequentially as blocks as a way of predicting school effectiveness (see Procedural Appendix for model).

The approach to selecting the final predictors incorporated both theoretical and "data driven" considerations. For example, efforts were made to explore the following questions:

- Once achievement has been adjusted for the effects of students' home background, what approaches to mathematics curriculum and instruction are associated with more effective schools?
- Are private schools more effective than public schools?
- Are public schools that replicate the characteristics of private schools just as effective?

- What is the relationship between factors that can be changed by schools and those that cannot?
- What is the relationship between students taking more advanced mathematics courses and other school effectiveness factors?
- Which background variables collected by NAEP are most highly related to variations in mathematics performance?

Blocks of variables were associated with each of these questions. The order of entry of the blocks into the model was designed to give descriptions of school processes as they actually occur. Thus, for example, the block reserved for mathematics instructional variables was entered early in the model to maximize the ability to detect these relationships before including in the analysis other predictor variables, such as those in the demographic block. Since the demographic variables are strongly correlated with general school characteristics, entering these variables first would reduce or eliminate relationships that could otherwise be identified. Priority also was placed on maintaining the significance of the variables at the point of their entry into the model and maximizing the proportion of variance explained by the various blocks. Thus, the blocks, the variables within blocks, and the block order differ somewhat from grade to grade. The increase in variance due to adding each block to the regression equations was tested for significance using an incremental procedure (see Procedural Appendix). Across the three grades, each block significantly increased the explained variance, between schools, with the exception of the second block at grade 4 — private/public school.

The results for grades 4, 8, and 12 are presented in Tables 2.2, 2.3, and 2.4, respectively.

Table 2.2

Predictors of School Effectiveness in Mathematics, Grade 4, 1992 NAEP

	BLOCK 1	BLOCK 2	BLOCK 3	BLOCK 4	BLOCK 5	ONLY
	Mathematics Curriculum/ Instruction	Private/ Public School	School Climate	Fewer Students Change Schools	Demographic Composition of Student Body	Demographic Composition of Student Body (Block 5)
	$R^2 = .34$	$R^2 = .35$	$R^2 = .44$	$R^2 = .53$	$R^2 = .80$	$R^2 = .73$
PREDICTORS	Regression Coefficient Estimates (with Probabilities indicating significance levels in Parentheses)					
Moderate Math Testing	**24.90 (.00)**	23.70 (.00)	21.49 (.00)	18.37 (.00)	9.35 (.00)	
Private versus Public School		**5.59 (.01)**	1.93 (.39)	−2.10 (.35)	−6.53 (.00)	
Positive Attitudes in School			**1.49 (.00)**	1.18 (.00)	0.28 (.38)	
Fewer Problems in School			**0.94 (.04)**	0.66 (.12)	0.16 (.64)	
Fewer Students Change Schools				**13.75 (.00)**	5.65 (.00)	
Higher Percentages White and Asian Students					**19.93 (.00)**	**26.08 (.00)**
School More Advantaged on SES					**2.29 (.00)**	**2.39 (.00)**

R^2 = Proportion of variance explained. As a guideline, $p < .05$ indicates a statistically significant predictor of school effectiveness in mathematics. T-tests have between imputation error added (see Appendix for information about imputations).

Note: Regression coefficients in bold-face type indicate the stage at which predictors were entered into the model.

Table 2.3

Predictors of School Effectiveness in Mathematics, Grade 8, 1992 NAEP

	BLOCK 1	BLOCK 2	BLOCK 3	BLOCK 4	BLOCK 5	BLOCK 6	ONLY
	Mathematics Curriculum/ Instruction	Private/ Public School	Students Plan to Take Geometry Grade 9	School Climate	Fewer Students Change Schools	Demographic Composition of Student Body	Demographic Composition of Student Body (Block 6)
	$R^2 = .40$	$R^2 = .42$	$R^2 = .58$	$R^2 = .63$	$R^2 = .67$	$R^2 = .85$	$R^2 = .73$
PREDICTORS	Regression Coefficient Estimates (with Probabilities indicating significance levels in Parentheses)						
Moderate Math Testing	**24.13 (.00)**	23.80 (.00)	20.42 (.00)	16.77 (.00)	14.22 (.00)	7.03 (.01)	
Frequently Use Calculators	**6.16 (.00)**	5.82 (.00)	4.25 (.00)	3.69 (.00)	3.75 (.00)	1.86 (.03)	
Frequently Do Math Problems	**18.16 (.00)**	16.22 (.00)	11.43 (.01)	11.40 (.01)	8.73 (.04)	3.30 (.30)	
Private versus Public School		**7.02 (.01)**	2.61 (.29)	−2.82 (.31)	−4.16 (.12)	−3.94 (.07)	
Students Plan to Take Geometry Grade 9			**22.24 (.00)**	20.77 (.00)	19.73 (.00)	13.35 (.00)	
Positive Attitudes in School				**.24 (.57)**	.41 (.31)	.71 (.02)	
Fewer Problems in School				**2. 26 (.07)**	2.41 (.04)	1.65 (.08)	
Less Impetus for Changes in School				**.85 (.00)**	.67 (.02)	0.001 (.99)	
Fewer Students Change Schools					**20. 73 (.00)**	13.97 (.00)	
Higher Percentages White and Asian Students					**20.52 (.00)**	**24.26 (.00)**	
School More Advantaged on SES						**1.34 (.00)**	**2.97 (.00)**

R^2 = Proportion of variance explained. As a guideline, $p < .05$ indicates a statistically significant predictor of school effectiveness in mathematics. T-tests have between imputation error added (see Appendix for information about imputations).

Note: Regression coefficients in bold-face type indicate the stage at which predictors were entered into the model.

Table 2.4

Predictors of School Effectiveness in Mathematics, Grade 12, 1992 NAEP

	BLOCK 1	BLOCK 2	BLOCK 3	BLOCK 4	BLOCK 5	BLOCK 6	ONLY	ONLY
	Mathematics Curriculum/ Instruction	Private/ Public School	School Climate	Students Took Algebra In or Before Grade 9	Demographic Composition of Student Body	Students Took More Advanced Math Courses	Demographic Composition of Student Body (Block 5)	Students Took More Advanced Math Courses (Block 6)
	$R^2 = .53$	$R^2 = .57$	$R^2 = .67$	$R^2 = .71$	$R^2 = .78$	$R^2 = .90$	$R^2 = .60$	$R^2 = .72$
PREDICTORS	Regression Coefficient Estimates (with Probabilities indicating significance levels in Parentheses)							
Moderate Math Testing	**19.68 (.00)**	16.42 (.00)	16.19 (.00)	16.74 (.00)	11.96 (.00)	10.43 (.00)		
Frequently Use Calculators	**11.63 (.00)**	11.14 (.00)	9.29 (.00)	8.11 (.00)	5.37 (.01)	4.46 (.01)		
Frequently Do Math Problems	**14.59 (.00)**	10.82 (.00)	8.62 (.01)	10.00 (.00)	9.55 (.00)	5.69 (.03)		
School Offers Advanced Math Courses	**3.73 (.00)**	3.68 (.00)	2.91 (.00)	2.40 (.00)	1.22 (.10)	−.50 (.45)		
Private versus Public School		**8.08 (.00)**	−1.39 (.61)	−2.79 (.29)	−3.77 (.13)	−5.13 (.01)		
Positive Attitudes in School			**9.37 (.00)**	7.41 (.00)	6.32 (.00)	3.43 (.06)		
Less Impetus for Changes			**1.95 (.02)**	1.54 (.06)	1.38 (.06)	−.09 (.89)		
Fewer Problems in School			**0.57 (.02)**	0.44 (.07)	0.12 (.61)	0.12 (.56)		
Students Took Algebra In or Before Grade 9				**20.08 (.00)**	15.50 (.00)	−5.84 (.21)		
Higher Percentages White and Asian Students					**10.47 (.00)**	13.45 (.00)	**13.32 (.00)**	
School More Advantaged on SES					**1.55 (.00)**	.85 (.03)	**3.82 (.00)**	
Student Took More Advanced Math Courses						**12.79 (.00)**		**16.39 (.00)**

R^2 = Proportion of variance explained. As a guideline, $p < .05$ indicates a statistically significant predictor of school effectiveness in mathematics. T-tests have between imputation error added (see Appendix for information about imputations).

Note: Regression coefficients in bold-face type indicate the stage at which predictors were entered into the model.

Mathematics Instruction and Curriculum

At each of the three grades, the first block was reserved for those NAEP background variables related to mathematics curriculum and instruction. Because factors related to curricular and instructional approaches are considered to be under the purview of schools, theoretically changes could be made in such practices to improve mathematics achievement. Entering these "alterable" variables into the model in the first block provided the best opportunity to examine the association between these factors and achievement in isolation from other variables.

Interestingly, among the many NAEP variables related to the delivery of school mathematics, the three strongest were:

- a moderate amount of testing
- doing problems from textbooks more often
- using calculators more frequently

At grade 4, the only variable found to be significant was moderation in the degree of testing in mathematics classes. At grade 4, this variable was defined as two categories based on students' reports that they took tests in mathematics once or twice a week or once or twice a month, rather than almost every day or hardly ever or never.

At grades 8 and 12, all three variables were found to be significant. The moderation in testing variable was categorized as taking mathematics tests once or twice a month, compared to once or twice a week, compared to almost every day or hardly ever. Common sense suggests that extreme amounts of testing (either daily or never) would not be particularly effective. However, the finding that at grades 8 and 12 monthly testing was more highly related to adjusted achievement than weekly testing is more difficult to interpret. It may be that teachers believe that better students need less constant monitoring, or that less testing leaves more time for instruction.

That solving more mathematics problems from textbooks and using calculators were associated with higher mathematics achievement is consistent with research in mathematics learning and recommendations contained in *The NCTM Standards* (see Chapter One). Both variables can be related to increasing time on task and providing students with more challenging problem situations. Doing mathematics problems from textbooks was coded into a frequency scale with three categories (almost every day, once or twice a week combined with once or twice a month, and never or hardly ever), while the four original response categories were kept for using calculators (almost every day, once or twice a week, once or twice a month, and never or hardly ever).

At grade 12, one more variable was found to be significant. High schools that offered a range of advanced mathematics courses were determined to be more effective. This variable was based on the weighted sum of schools' reports that they taught courses at least one semester in length in the following subjects: trigonometry; pre-calculus and/or Algebra III and/or elementary functions; probability and/or statistics; and calculus in combination with their reports about the percentages of students enrolled in AP calculus (>50, 26 to 50, 11 to 25, 6 to 10, 1 to 5, and 0).

Although the variables generally associated with effective mathematics instruction were highly related to each other and the number that eventually emerged as significant in this analysis was small, the first block accounted for 34 percent of the variance at grade 4, 40 percent at grade 8, and 53 percent at grade 12.

Private Versus Public Schools

The second block entered at all three grades consisted of a single variable — whether the school was a Catholic or other type of private school, as differentiated from a public school. Obviously, this variable was not entered in the second position because it can be altered. The purpose here was to explore the effectiveness of private schools compared to their public school counterparts adjusting for the effects of students' home background and schools' socioeconomic environment. Also, including factors generally associated with private schools (e.g., fewer problems, more students taking advanced courses, and greater parental interest) in subsequent blocks enabled some examination of these factors vis-à-vis the effectiveness of private as compared to public schools.[44]

As shown in Tables 2.2 through 2.4, the amount of additional variance explained by this second block was small, 1 to 4 percent across the three grades. Also, in combination with the blocks of school climate variables that were added subsequently, the effects of private versus public school were no longer significant at that stage. In essence, these results suggest that the effectiveness of private schools can be replicated if students, teachers, and

[44] Raudenbush, S. and Bryk, A., "A Hierarchical Model for Studying School Effects," *Sociology of Education*, 59, 1-17, 1986.

parents have positive attitudes toward academic achievement, the composition of the student body is stable, larger proportions of students are taking advanced courses, and problems are few, as denoted by several of the next blocks.

School Climate

Three composite variables or indices were included in the block measuring school climate or environment. These were:

- positive attitudes toward academic achievement
- absence of problems in the school
- limited impetus for change (at grades 8 and 12)

At all three grades, school administrators were asked to characterize the prevailing climate regarding factors related to support for academic achievement in each of their schools. The scale used was: very positive, somewhat positive, somewhat negative, or negative. In particular, the composite variable developed for this analysis was based on the sum of their responses to questions about student attitudes toward academic achievement, parental support for student achievement, relations between students and teachers, teacher morale, and regard for school property.

Similarly, school administrators were asked about a variety of problems found in schools and asked to categorize each as: not a problem, minor, moderate, or serious. At grade 4, responses were summed for the following: student absenteeism, physical conflicts among students, teacher absenteeism, racial or cultural conflict, and student health problems. At grades 8 and 12, student tardiness, student cutting of classes, and the percentage of students absent on a given day (0 to 2, 3 to 5, 6 to 10, and more than 10 percent) also were included in the school problems indicator.

At grades 8 and 12, a third composite variable was included in this block, based on school administrators' responses about the extent to which various sources had served as an impetus to change curriculum or instructional practices. Responses were coded dichotomously ("not at all" versus to "some" or "a great extent") and summed across the following two sources at grade 8: 1) district or school testing programs and 2) state mandates. At grade 12, two additional sources were included: 1) public reporting of school or district performance data and 2) budget data.

Across the three grades, this block accounted for an additional 5 to 10 percent of the variance. Please note that it was entered in the fourth position at grade 8, after students' reports about their plans for ninth-grade mathematics course taking. That may have reduced its contribution at that grade and also may account for the fact that the composite variable for positive school climate or attitudes was not significant at grade 8, although it was at grades 4 and 12. As noted in Chapter One, the "character" of schools, as signified by a general focus or press toward academic achievement, has been found to be very important in fostering higher achievement. A school climate of positive attitudes toward academic achievement was found to be significant at grade 8 after adjusting for the socioeconomic level of the school. The school problems composite was moderately significant when entered and in the presence of other blocks, except at grade 12 when the school socioeconomic block was entered. The composite variable based on various sources serving as an impetus for change, although difficult to interpret, also was moderately significant throughout the analysis, except at grade 12 in the presence of the mathematics course taking block.

Stability of the Student Body

At grades 4 and 8, the block following school climate (either block four or five, respectively) contained results to a single question asked of students about how many times they had changed schools in the past two years because they had changed where they lived. The importance of this variable is consistent with considerable research dating back across several decades regarding positive relationships between the stability of the student body and achievement.[45] In general, a high turnover of students is thought to disrupt continuity in teaching. Even entered subsequent to a number of rather strong predictors, this variable accounted for an additional 9 percent of the variance between effective schools at grade 4 and an additional 4 percent at grade 8. At grade 12, however, student body stability was not a measured variable.

[45] Bryant, E. C., Glaser, E., Hansen, M. H., and Kirsch, A., *Associations between Educational Outcomes and Background Variables: A Review of Selected Literature* (Denver, CO: The National Assessment of Educational Progress, 1974).

Tracking and Taking Advanced Mathematics Courses at Earlier Grades

It may seem evident that taking mathematics courses would increase academic achievement and that course taking would matter despite home background, school climate, and other factors. Both the amount and quality of instructional time have emerged consistently as powerful factors in the syntheses of thousands of educational research results.[46] Nevertheless, as emphasized in Chapter One, the majority of students in the United States do not take advanced courses in mathematics. The mathematics curriculum in our country tends to be sharply differentiated by track. College-bound students take a markedly different sequence of courses than do students in vocational/technical or general high school programs. Among those in academic programs, honor students often take an even more advanced sequence. Regardless of track, however, the prescribed sequence of courses is generally consistent — pre-algebra, Algebra I, and geometry before Algebra II and trigonometry, and all those before Algebra III or pre-calculus, and then on to calculus.

It then follows that students who begin taking algebra at an earlier grade have been deemed as being in a higher academic track, and therefore, from one perspective, schools with more students taking algebra in early grades may simply have more able students than their counterparts. On the other hand, research has shown that the practice of such grouping for mathematics instruction also can curtail opportunities to learn for students in the non-academic tracks.[47] At the most obvious level, a late entry into the course-taking sequence limits options for the number of courses that can be taken prior to high school graduation. Perhaps less obvious are differences in the instructional approaches used in advanced courses compared to less advanced courses. For example, NAEP data indicate more use of calculators in advanced courses.[48] There also are indications that teachers in advanced courses are more likely to emphasize higher-order thinking skills.[49]

---

[46] Walberg, H.J., "Productive Time and Subject Matter Learning." In D. Tanner and J.W. Keefe, editors, *Improving the Curriculum: The Principal's Challenge* (Reston, VA: National Association of Secondary School Principals. 1989).

[47] Oakes, J., *Multiplying Inequalities: The Effects of Race, Social Class, and Tracking on Opportunities to Learn Mathematics and Science* (Santa Monica, CA: The Rand Corporation, 1990).

[48] Dossey, J.A., Mullis, I.V.S., & Gorman, S., *How School Mathematics Functions* (Washington, DC: National Center for Education Statistics, 1994).

[49] Raudenbush, S. W., Rowan, B., and Cheong, Y. F., "Higher Order Instructional Goals in Secondary Schools: Class, Teacher, and School Influences," *American Educational Research Journal*, Vol.30(3), 1993.

The third block in the grade 8 analysis contained students' reports about the mathematics class they expected to take in the ninth grade. The variable was defined as three categories: one category for those eighth graders expecting to take geometry, a second category for those expecting to take Algebra I, and a third category for students planning to take other courses (e.g., general mathematics, pre-algebra, consumer mathematics) or those who did not know their plans. Because of the course-taking sequences previously noted, schools having a greater percentage of students expecting to take geometry in the ninth grade most likely had greater percentages of eighth graders enrolled in Algebra I and more seventh graders in pre-algebra, etc. Ability grouping in mathematics begins early, so that by ninth grade, the percentage of students prepared to take geometry classes already is severely limited, as is, to some extent, the percentage prepared to take Algebra I classes (see Chapter One). This block accounted for an additional 16 percent of the variance at grade 8, when entered following the mathematics instructional variables and private versus public school. The percentages of students planning to take higher-level mathematics courses in the ninth grade remained a highly significant variable throughout the remainder of the blockwise analysis at grade 8.

At grade 12, the indicator of academic track was based on students' responses to a question asking in which grade they first took Algebra I. Their responses were coded dichotomously: having taken Algebra I in or before grade 9 versus having taken Algebra I after grade 9 or not at all. When entered subsequent to school climate (one block later than at grade 8), the tracking variable accounted for somewhat less additional variation among schools (4 percent). However, it was easily significant.

Demographic Composition of the School's Student Body

The last block entered at grades 4 and 8, and the second to the last block at grade 12, contained two variables — the first was based on students' report of their race/ethnicity and the second was a complicated index of various school socioeconomic indicators. For the race/ethnicity variable, students classifying themselves as of White or Asian/Pacific Islander background were in one category and those reporting African American or Hispanic heritage were in the second category. (American Indian students were not included in the analysis, because of their very small sample size.) The school socioeconomic index had three components: 1) the students' home background variable aggregated to the school level (based on students' reports of reading materials in the home, highest level of parents' education, and number of parents or stepparents living at home as described for the first HLM analysis discussed in this chapter); 2) the percentages of students receiving the subsidized lunch and/or nutrition program (0 to 5, 6 to 25, and 26 to 100 percent); and 3) the size and type of community in which the school was located (some adjustments were made grade by grade, but essentially four community types were coded, ranging from advantaged urban areas to disadvantaged urban areas, also as explained earlier in this chapter).

The strength of this block was apparent at all three grades. At grade 4, the demographic composition of the student body accounted for an additional 27 percent of the variation across schools, and when entered as the sole block in the model, accounted for 73 percent of the variation in school effectiveness. Similarly, at grade 8, this block accounted for an additional 18 percent of the variation and 73 percent when entered all by itself. At grade 12, differences in the demographic characteristics of the student body only accounted for an additional 7 percent of the variation in achievement, although singly this block explained 60 percent of the variation in effectiveness, and the two variables were significant.

Clearly, although the initial blocks at each grade accounted for a substantial portion of the outcome variance, much of this prediction was correlated with the home background and demographic characteristics of the student body. In particular, at grades 4 and 8, a substantial portion of the school demographics block's predictive strength was independent of the previous blocks (27 and 18 percent, respectively). This block, in part, also may represent the effects of important variables that were not collected by NAEP, but were related to general school and community characteristics. NAEP, like most assessment efforts, is unavoidably incomplete, lacking adequate information about some important variables and no information about others.[50]

Taking Advanced Courses in High School

At grade 12, one last single-variable block was entered into the model — twelfth graders' reports on the number of advanced mathematics courses they had taken in high school. Five courses were included: geometry;

[50] Koretz, D., *Evaluating and Validating Indicators of Mathematics and Science Education* (Santa Monica, CA: The Rand Corporation, 1992).

Algebra II; trigonometry; pre-calculus/Algebra III/elementary functions; and calculus. Responses were coded as a sum of at least some study (based on options — one-half year or less, one school year, or more than one year) versus no study of each of the subjects.

While effects of tracking, as previously noted, cloud the interpretation of these results, the data support an interpretation that mathematics course taking can account for variation in effectiveness above and beyond that of the home background and demographic characteristics of the student body. This block accounted for an additional 12 percent of the variation, when entered as the sixth block. When entered singly, this block accounted for 72 percent of the variation. Schooling did make a significant difference in mathematics achievement beyond the effects of home background and community, particularly when students pursued rigorous mathematics coursework. Unfortunately, among other things, as described throughout this report, tracking and course sequencing practices, combined with students' dwindling interest in mathematics as they progress through school, tend to limit mathematics course taking.

Correlation Matrices

The correlations of the school-level predictors are presented in Tables 2.5 through 2.7 for grades 4, 8, and 12, respectively. At grade 4, the correlations range from .18 to .47 among predictors in the first four blocks, indicating weak to moderately large associations among these factors. The correlations among predictors in the first four blocks and the two predictors in the demographic characteristics block — race/ethnicity and school socioeconomic indicators — ranged from .28 to .52, further indicating the overlap between the demographic characteristics of the school's student body and predictors in the first four blocks.

Table 2.5

Correlations of School-Level Predictors, Grade 4

	Moderate Math Testing	Private School	Positive School Attitudes	Fewer School Problems	Fewer Change Schools	Higher Percent White and Asian	Advantaged School SES
Moderate Math Testing	1.00	.18	.18	.20	.31	.45	.37
Private School	.18	1.00	.25	.38	.43	.28	.52
Positive School Attitudes	.18	.25	1.00	.47	.32	.35	.44
Fewer School Problems	.20	.38	.47	1.00	.34	.35	.43
Fewer Change Schools	.31	.43	.32	.34	1.00	.49	.52
Higher % White and Asian Students	.45	.28	.35	.35	.49	1.00	.62
Advantaged School SES	.37	.52	.44	.43	.52	.62	1.00

Table 2.6

Correlations of School-Level Predictors, Grade 8

	Moderate Math Testing	Do More Math Problems	More Use Calculators	Private School	Plane Geometry Grade 9	Positive School Attitudes	Less Impetus Change	Fewer School Problems	Fewer Change Schools	Higher % White and Asian Students	Advantaged School SES
Moderate Math Testing	1.00	.20	.25	.09	.22	.13	.17	.30	.28	.45	.34
Do More Math Problems	.20	1.00	.00	.17	.21	.13	.08	.12	.23	.28	.25
More Use Calculators	.25	.00	1.00	.12	.24	.22	.10	.23	.09	.29	.35
Private School	.09	.17	.12	1.00	.30	.39	.30	.48	.25	.18	.45
Plane Geometry Grade 9	.22	.21	.24	.30	1.00	.35	.14	.27	.21	.35	.51
Positive School Attitudes	.13	.13	.22	.39	.35	1.00	.13	.56	.11	.18	.37
Less Impetus Change	.17	.08	.10	.30	.14	.13	1.00	.23	.11	.18	.29
Fewer School Problems	.30	.12	.23	.48	.27	.56	.23	1.00	.29	.40	.53
Fewer Change Schools	.28	.23	.09	.25	.21	.11	.11	.29	1.00	.30	.43
Higher % White and Asian Students	.45	.28	.29	.18	.35	.18	.18	.40	.30	1.00	.67
Advantaged School SES	.34	.25	.35	.45	.51	.37	.29	.53	.43	.67	1.00

Table 2.7

Correlations of School-Level Predictors, Grade 12

	Moderate Math Testing	More Use Calculators	Do More Mathematics Problems	School Offers Advanced Math	Private School	Positive School Attitudes	Less Impetus Change	Fewer School Problems	Took Algebra Before/In Grade 9	Higher Percents White and Asian	Advantaged School SES	Students Took Advanced Math
Moderate Math Testing	1.00	.02	.45	.01	.06	.07	.03	.03	.03	.25	.12	.01
More Use Calculators	.02	1.00	.53	.19	.28	.24	.34	.33	.31	.37	.43	.44
Do More Math Problems	.45	.53	1.00	.24	.32	.33	.31	.26	.21	.08	.36	.45
School Offers Advanced Math	.01	.19	.24	1.00	.12	.22	.18	.13	.25	.24	.40	.45
Private School	.06	.28	.32	.12	1.00	.42	.68	.51	.44	.22	.51	.57
Positive School Attitudes	.07	.24	.33	.22	.42	1.00	.25	.52	.43	.22	.47	.53
Less Impetus Change	.03	.34	.31	.18	.68	.25	1.00	.42	.41	.25	.45	.59
Fewer School Problems	.03	.33	.26	.13	.51	.52	.42	1.00	.45	.37	.53	.50
Took Algebra Before/In Grade 9	.03	.31	.21	.25	.44	.43	.41	.45	1.00	.35	.49	.70
Higher % White and Asian Students	.25	.37	.08	.24	.22	.22	.25	.37	.35	1.00	.58	.33
Advantaged School SES	.12	.43	.36	.40	.51	.47	.45	.53	.49	.58	1.00	.64
Students Took Advanced Math	.01	.44	.45	.45	.57	.53	.59	.50	.70	.33	.64	1.00

Similarly, at grade 8, with the exception of the association between positive attitudes and fewer school problems (.56) and between fewer problems and private schools (.48) the correlations tended to be smaller among predictors in the first five blocks than between the two predictors in the demographics block (.67) or between those variables in the demographics block and other predictor variables. Particularly notable were the relationships between problems in the school and the school socioeconomic indicators (.53), as well as between tracking (geometry in ninth grade) and the school socioeconomic indicators (.51).

At grade 12, the pattern continued, although the relationships among moderate testing and doing more problems (.45) as well as between doing problems and using calculators (.53) are interesting. Of most interest at grade 12 are the correlations between taking advanced courses and the remaining predictors in the model. With the exception of moderate testing in mathematics class, these ranged from .33 through .70. The strongest association, as might be anticipated, was between tracking/getting an early start in advanced coursework and the number of advanced courses taken during high school.

Summary

Because learning is an outcome of both in-school and out-of-school contexts and experiences, it is important to examine mathematics achievement in the larger social context beyond schooling. It is also important to consider the large impact that students' home background has on achievement and to explore which school factors have an effect beyond that of the home and community. From the perspective of coping most effectively given the difficulty of their educational task, the most effective schools are not necessarily those with the highest achieving students, but the ones that make the most difference beyond the effects of home and community.

Because the NAEP data are hierarchical in structure — students within teachers' classes, classes within schools, and schools within states — it is appropriate to take advantage of new technology designed to provide more precise estimates based on nested data in analyzing the NAEP data for information about effective schools. By first conducting rather broad-based HLM analyses, two important goals were reached. First, an array of NAEP variables could be brought to bear in an analysis of school effectiveness. Second, much has been learned methodologically — both about associations among the NAEP background variables and the difficulties that might be encountered in conducting hierarchical analyses with complex samples such as those in the NAEP database.

Two different analyses of effective schools were conducted, each involving two levels — students within schools and schools themselves.[51] Approximately 5,000 students and 200 schools were involved in each of the analyses at each of the three grades.

[51] Jenkins, F., *Using Hierarchical Analysis to Identify Factors Contributing to Effective Schools*, Research Report, Educational Testing Service, (in progress).

In the first analysis, average mathematics proficiency was adjusted within each school based on the home background of the individual students (reading materials in home, parents' education level, and number of parents or stepparents in home). Then, the most and least effective schools were identified based on whether or not they exceeded the achievement typical of schools with similar socioeconomic characteristics (percentages of students in subsidized lunch programs, the schools' community type, and the home backgrounds of students in the school). Comparisons between the most and least effective schools with respect to the NAEP background variables showed a number of significant differences. At all three grades, the most effective schools had students who watched less television, students who changed schools less often, and students who were subjected to only a moderate amount of testing in their mathematics classes (about weekly to monthly). At grade 8 and particularly at grade 12, matters of more students in advanced tracks, students taking more advanced courses, positive attitudes toward academics, and fewer problems in the schools also became distinguishing features of the most effective schools. For example, at grades 8 and 12, both the students and their parents had more positive attitudes toward academic achievement. In the most effective high schools, some additional differences related to a focus on academic achievement were noted. At grade 12, more pages were read for school each day, more mathematics problems were worked from textbooks on a daily basis, and more students went on to four-year colleges.

The second HLM analysis used average school achievement adjusted for students' home background as the measure of effectiveness. A blockwise analysis was then conducted to identify sets of variables associated with more effective schools. These analyses tended to confirm numerous previous analyses indicating the large predictive power of socioeconomic factors. Yet, variables related to mathematics curriculum and instruction and school climate were able to account for substantial amounts of variation in achievement across schools. Also, students' having taken more advanced mathematics courses was the single most powerful predictor of higher mathematics achievement at grade 12.

More particularly, moderate testing in mathematics classes (approximately monthly), doing mathematics problems more frequently, and more frequent use of calculators appeared to be strong factors. The mathematics instruction block was entered first and accounted for about one-third of the variance at grade 4, about 40 percent at grade 8, and about half at grade 12. Blocks related to school atmosphere, including the stability of the student body and more students in advanced courses, accounted for an additional 20 percent or so of the variance across the grades.

The results indicated that the effectiveness of private schools overlaps with that of public schools in which students, teachers, and parents have positive attitudes toward academics, and few problems exist (for example, regarding physical conflicts or absenteeism). At grades 4 and 8, a more stable student body with students who changed schools fewer times was associated with higher school effectiveness. At grades 8 and 12, greater percentages of students taking advanced mathematics courses (algebra and geometry in grades 8 and 9) predicted higher effectiveness. Of course, the practice of tracking more able students into more advanced mathematics courses makes this a complex finding to interpret. Blocks related to school atmosphere, including the stability of the student body and more students in advanced courses, accounted for an additional 20 percent or so of the variance across the grades.

Adding an index of several school-level socioeconomic indicators accounted for an additional 20 to 25 percent of the variance at grades 4 and 8, and about 7 percent at grade 12. Thus, the complete model accounted for at least 80 percent of the variance at all three grades. However, only using the school-level demographic index accounted for about 60 to 70 percent of the variance.

Mathematics course taking, as a single-variable block, was added as the very last block at grade 12. More advanced mathematics course taking (geometry, advanced algebra, trigonometry, and calculus) was able to account for an additional 12 percent of the variance, after taking into account the effects of students' home background, mathematics instruction, school atmosphere, tracking, and school demographic indicators. Entered as the only block in the analysis, taking advanced courses was able to account for 72 percent of the achievement variation at grade 12. As might be expected, however, this variable was highly correlated with other variables in the model.

In summary, students' home background and school socioeconomic indicators are powerful influences on academic achievement in mathematics. However, schools can make a difference beyond the impact of home and community. In particular, if a school's atmosphere encourages learning, and students take more challenging mathematics courses, the school is likely to be more effective.

Procedural Appendix

Overview of Procedures and Methods

The 1992 Mathematics Assessment Framework and Questions

The framework underlying NAEP's 1992 mathematics assessment was initially developed for the 1990 assessment and subsequently approved for use in both assessments by the National Assessment Governing Board. It was developed through a consensus process managed by the Council of Chief State School Officers, and the items were developed through a similarly broad-based process managed by Educational Testing Service. The development of the 1992 mathematics assessments benefited from the involvement of hundreds of representatives from State Education Agencies who attended numerous NAEP NETWORK meetings; served on committees; reviewed the framework, objectives, and questions; and in general, provided important suggestions on all aspects of the program.

The mathematics assessment framework is a five-by-three matrix specifying five content areas — Numbers and Operations; Measurement; Geometry; Data Analysis, Statistics, and Probability; and Algebra and

Functions; plus three process or ability areas. These include Conceptual Understanding, Procedural Knowledge, and Problem Solving.[52] Consistent with standards developed by the National Council of Teachers of Mathematics, many questions required students to construct their responses, and some questions asked for explanations of their reasoning.[53] For various portions of the assessment, mathematical tools and aids were supplied, including scientific calculators, protractor/rulers, and geometric shapes. The analyses discussed in this report were based on 155 questions at grade 4, 183 questions at grade 8, and 179 questions at grade 12.

Each student received a booklet containing a set of general background questions, a set of subject-specific background questions, three 15-minute segments, or blocks, of cognitive items, and a set of questions about his or her motivation and familiarity with the assessment material. At each grade level, the achievement measure used for the analyses in this report was based on 13 different blocks of multiple-choice and constructed-response content questions. Students received different blocks of cognitive items in their booklets according to a careful plan. The 1992 assessment was based on an adaptation of matrix sampling called balanced incomplete block (BIB) spiraling — a design that enables broad coverage of mathematics content while minimizing the burden for any one student. The balanced incomplete block part of the design assigns blocks of items to booklets and each pair of blocks appears together in at least one booklet. The spiraling part of the method cycles the booklets for administration, so that typically only a few students in any assessment session receive the same booklet.

In accordance with this design, the 13 blocks were presented in 26 booklets. Each block appeared in exactly six booklets, and each block appeared with every other block in at least one booklet. Students at grades 4 and 8 were given calculators to use with three of the 13 blocks and were trained in their use prior to the assessment. Students at grade 12 were given calculators to use with four of the 13 blocks. At the fourth grade, students were provided with four-function calculators and at grades 8 and 12, they were provided with scientific calculators. For another block, fourth-grade students were provided with a ruler, and eighth- and twelfth-grade students

[52] *Mathematics Objectives, 1990 Assessment* (Princeton, NJ: National Assessment of Educational Progress, Educational Testing Service, 1988).

[53] Dossey, J. A., Mullis, I. V. S., & Jones, C. O., *Can Students Do Mathematical Problem Solving? Results from Constructed Response Questions in NAEP's 1992 Mathematics Assessment* (Washington, DC: National Center for Education Statistics, Government Printing Office, 1993).

with a protractor/ruler. For still another block, at all three grades, students were given geometric shapes (manipulatives) to provide a concrete basis for determining their answers.

Background Questionnaires

As part of the 1992 mathematics assessment, students, teachers, and school administrators completed background questionnaires. The analyses in this report involve data collected from students and from the principal (or another designated administrator) in each participating school. The data collected from the mathematics teachers of the fourth- and eighth-grade students participating in the 1992 assessment were not included because the small number of mathematics teachers per grade/per school complicates multilevel analyses (see section describing the hierarchical analyses).

An expert panel knowledgeable about NAEP, educational policy, and instruction in the curriculum areas being assessed in 1992 developed guidelines for the student, school, and teacher questionnaires. The framework focused on five educational areas: instructional content, instructional practices and experiences, teacher characteristics, school conditions and contexts, and conditions beyond school (i.e., home support, out-of-school activities, and attitudes).[54] The outline for the background questionnaire framework follows.

NAEP 1992 BACKGROUND QUESTIONNAIRE FRAMEWORK

1.0 INSTRUCTIONAL CONTENT

- 1.1 course offerings in selected subject areas
- 1.2 course taking in selected subject areas
- 1.3 objectives, topics, and skills covered
- 1.4 emphasis on facts, concepts, and higher-order skills

2.0 INSTRUCTIONAL PRACTICES AND EXPERIENCES

- 2.1 assignment to classes according to ability or achievement
- 2.2 grouping within classes according to ability or achievement

[54] *National Assessment of Educational Progress, 1992 Policy Information Framework* (Princeton, NJ: National Assessment of Educational Progress, Educational Testing Service, 1992).

2.3 teacher's freedom within the classroom

2.4 time spent on content-related instruction

2.5 use of whole class, group, and individual instruction

2.6 mode of instruction — lecture, demonstration, discussion, etc.

2.7 availability and use of materials — textbooks, supplementary materials, workbooks, kits

2.8 availability and use of equipment — computers and calculators

2.9 classroom activities and assignments

2.10 student assessment

2.11 amount of homework assigned

3.0 TEACHER CHARACTERISTICS

3.1 type of certification

3.2 highest academic degree

3.3 undergraduate and graduate coursework in mathematics, reading or writing, and in the teaching of those subjects

3.4 undergraduate and graduate major and minor field

3.5 in-service training in mathematics, reading or writing, and in the teaching of those subjects

3.6 other teacher development activities

3.7 number of years teaching experience in general

3.8 number of years teaching in a field

3.9 comfort in teaching mathematics

4.0 SCHOOL CONDITIONS AND CONTEXT

4.1 instructional time and teacher-pupil ratio

4.2 schoolwide programs

4.3 characteristics and experience of the principal

4.4 characteristics and experience of the teaching staff

4.5 school climate

4.6 resources for students with special needs

4.7 resources for teachers

4.8 community and parental involvement

5.0 CONDITIONS BEYOND SCHOOL

- 5.1 language in home
- 5.2 country of birth
- 5.3 student mobility
- 5.4 home resources
- 5.5 parental support
- 5.6 experiences before starting school
- 5.7 out-of-school activities — reading
- 5.8 computer use
- 5.9 disposition to learning — attitudes toward subjects, self-confidence in subjects, value and utility of subjects, educational and vocational aspirations

It is important to note that in this report, as in all NAEP reports, the student is always the unit of analysis, even when information from the school questionnaire is being reported. Using the student as the unit of analysis makes it possible to describe the school conditions for representative samples of students. Although this approach may provide a different perspective from that obtained by simply reporting information collected from the principals, it is consistent with NAEP's goal of providing information about the educational context and performance of students.

National Sampling

Sampling and data collection activities for the 1992 NAEP assessment were conducted by Westat, Inc. In 1992, the national assessment was conducted from January through March, with some make-up sessions in early April.

As with all NAEP national assessments, the results for the national samples were based on a stratified, three-stage sampling plan. The first stage included defining geographic primary sampling units (PSUs), which are typically groups of contiguous counties, but sometimes a single county; classifying the PSUs into strata defined by region and community type; and randomly selecting PSUs. For each grade, the second stage included listing, classifying, and randomly selecting schools, both public and private, within each PSU selected at the first stage. The third stage involved randomly selecting students within a school for participation. Some students who were selected (about 7 to 8 percent) were excluded because of limited English proficiency or severe disability.

Table A.1 presents the student and school sample sizes and the cooperation and response rates for the national assessment.

Table A.1

1992 Student and School Sample Sizes

	Number of Participating Schools	Percent of Schools Participating	Number of Students	Percent of Student Completion
Grade				
4	527	86	8,738	93
8	587	84	9,432	89
12	468	81	8,499	81
Total	1,582		26,669	

Although sampled schools that refused to participate were occasionally replaced, school cooperation rates were computed based on the schools originally selected for participation in the assessments. The rates, which are based on schools sampled for all subjects assessed in 1992 (reading, writing, and mathematics), are also the best estimates for the mathematics assessment. The student completion rates represent the percentage of students assessed of those invited to be assessed in mathematics, including those assessed in follow-up sessions, when necessary.

Note: In 1992, NAEP also conducted a voluntary Trial State Assessment Program in mathematics at grades 4 and 8. Data for the 44 jurisdictions participating in the program can be found in the *NAEP 1992 Report Card for the Nation and the States*, which provides overall achievement results for various demographic subgroups and *Can Students Do Mathematical Problem Solving?*, which looks at the results for questions where students were asked to construct their responses. A relatively complete set of achievement and background data can be found in the *Data Compendium for the NAEP 1992 Mathematics Assessment of the Nation and the States.*

Excluded Students

It is NAEP's intent to assess all selected students. Therefore, all selected students who are capable of participating in the assessment should be assessed. However, some students sampled for participation in NAEP are excluded from the sample according to carefully defined criteria. Specifically, some of the students identified as having Limited English Proficiency (LEP) or having an Individualized Education Plan (IEP) may be incapable of participating meaningfully in the assessment. These students are identified as follows:

LEP students may be excluded if:

- the student is a native speaker of a language other than English, AND
- he or she has been enrolled in an English-speaking school for less than two years, AND
- the student is judged to be incapable of taking part in the assessment

IEP students may be excluded if:

- the student is mainstreamed less than 50 percent of the time in academic subjects and is judged to be incapable of taking part in the assessment, OR
- the IEP team has determined that the student is incapable of taking part meaningfully in the assessment

When there is doubt, the student is included in the assessment.

For each student excluded from the assessment, school personnel completed a questionnaire about the characteristics of that student and the reason for exclusion. Approximately 7 to 8 percent of the students nationally were excluded from the 1992 assessment.

Data Collection and Scoring

As with all NAEP assessments, data collection for the 1992 assessment was conducted by a trained field staff. For the national assessment, this was accomplished by Westat staff.

Materials collected as part of the 1992 assessment were shipped to National Computer Systems in Iowa City for processing. Receipt and quality control were managed through a sophisticated bar-coding and tracking system. After all appropriate materials were received from a school, they were forwarded to the professional scoring area, where the responses to the open-ended items were evaluated by trained staff using guidelines prepared by NAEP. Each open-ended question had a unique scoring guide that defined the criteria to be used in evaluating students' responses. Of the regular constructed-response items, most were scored right/wrong, but some included several different categories of correct and incorrect responses. The extended constructed-response questions were evaluated on a scale of 1 to 5, permitting degrees of partial credit to be given.

For the national mathematics assessment and the Trial State Assessment Program, approximately 4 million student responses were scored, including a 20 percent reliability sample. The overall percentage of agreement between readers for both the national and Trial State Assessment Program reliability samples at each of the three grades assessed was 94 percent. Subsequent to the professional scoring, the booklets were scanned, and all information was transcribed to the NAEP database at ETS. Each processing activity was conducted with rigorous quality control.

Data Analysis and IRT Scaling

After the assessment information had been compiled in the database, the data were weighted according to the population structure. The weighting for the national and state samples reflected the probability of selection for each student as a result of the sampling design, adjusted for nonresponse. Through poststratification, the weighting ensured that the representation of certain subpopulations corresponded to figures from the U.S. Census and the Current Population Survey.[55]

Analyses were then conducted to determine the percentages of students who gave various responses to each cognitive and background question. Item response theory (IRT) was used to estimate average scale-score proficiency for the nation and for various subgroups of interest within the nation. It is also used to determine the top-performing and bottom-performing one-third schools. IRT models the probability of answering an item correctly as a mathematical function of proficiency or skill. The main purpose of IRT analysis is to provide a common scale on which performance can be compared across groups, such as those defined by grades, and subgroups, such as those defined by race/ethnicity or gender. Because

[55] For additional information about the use of weighting procedures in NAEP, see Eugene G. Johnson, "Considerations and Techniques for the Analysis of NAEP Data" in *Journal of Educational Statistics* (December 1989).

of the BIB-spiraling design used by NAEP, students do not receive enough questions about a specific topic to provide reliable information about individual performance. Traditional test scores for individual students, even those based on IRT, would lead to misleading estimates of population characteristics, such as subgroup means and percentages of students at or above a certain proficiency level. Instead, NAEP constructs sets of plausible values designed to represent the distribution of proficiency in the population. A 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.[56]

To examine the relationship between the level of school performance and characteristics of students, schools, and teachers, NAEP sorted the schools by their students' average performance on the mathematics assessments (as measured by the average of the plausible values across all assessed students in the grade in the school). The top one-third and bottom one-third of the schools were then identified and the characteristics of these schools were then compared on a wide array of variables. The fact that a relationship exists between achievement and another variable, however, does not reveal the underlying cause of the relationship, which may be influenced by a number of other variables. Similarly, the assessments do not capture the influence of unmeasured variables. The results are most useful when they are considered in combination with other knowledge about the student population and the educational system, such as trends in instruction, changes in the school-age population, and societal demands and expectations.

Hierarchical Analyses

The Sample. The hierarchical analyses reported herein were based on the student and school data collected as part of NAEP's 1992 mathematics assessments at grades 4, 8, and 12. Because of the additional complexities involved, the teacher data and voluminous state-by-state data (both

[56] For theoretical justification of the procedures employed, see Robert J. Mislevy, "Randomization-Based Inferences About Latent Variables from Complex Samples," *Psychometrika*, 56(2), 177-196, 1988).

For computational details, see *Focusing the New Design: NAEP 1988 Technical Report* (Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress, 1990) and the *1990 NAEP Technical Report*.

available only at grades 4 and 8) were not included. After applying sample weights, the NAEP sample from the BIB-spiraled portions of the 1992 mathematics assessments originally consisted of 6,646 students in 380 schools at grade 4, 7,040 students in 345 schools at grade 8, and 6,661 students in 316 schools at grade 12. However, certain subgroups had very small sample sizes, including American Indian students and students in Department of Defense schools. The Hierarchical Linear Model (HLM) program required that there be no missing school data.[57] Also, since the procedure involved fitting a linear regression within each school to adjust for students' home background, a minimum sample size of 15 students was required to ensure adequate stability. Therefore, schools with missing school data or with fewer than 15 students were deleted, as were American Indian students and those in Department of Defense Schools. The final sample sizes for the hierarchical analyses are shown in Table A.2. Since the sampling weights for the original samples no longer applied to the subsamples used for the hierarchical analyses, the analyses were conducted using an unweighted sample. Table A.3 contains descriptive statistics comparing the weighted national samples and the unweighted subsamples used in these analyses.

Table A.2

Sample Sizes for Hierarchical Analyses

	GRADE 4	GRADE 8	GRADE 12
Students	5,081	4,979	4,905
Schools	224	186	189
Average Number Students Per School	23	27	26
Range Students Per School	15 to 55	15 to 64	15 to 52

[57] Bryk, A., Raudenbush, S., Seltzer, M. and Congdon, R., *An Introduction to HLM: Computer Program and User's Guide* (University of Chicago, Dept. of Education, 1989).

Table A.3

Descriptive Statistics for Original Weighted Samples and Unweighted Hierarchical Analysis Subsamples

	GRADE 4		GRADE 8		GRADE 12	
	Original Weighted	Subsample Unweighted	Original Weighted	Subsample Unweighted	Original Weighted	Subsample Unweighted
Number of Students	6,646	5,081	7,040	4,979	6,661	4,905
Average Proficiency	218	217	268	268	299	302
Standard Deviation	31	31	35	36	33	34
Percent White	70	64	70	66	71	71
Percent Black	16	16	16	16	15	15
Percent Hispanic	10	17	10	16	10	9
Percent With Parents' Education High School Graduate or Less	52	53	40	40	31	28
Percent With Parents' Education Beyond High School Graduation	48	47	60	61	69	71
Percent Extreme Rural	12	11	9	8	10	10
Percent Disadvantaged Urban	9	10	9	11	12	11
Percent Advantaged Urban	12	12	10	10	12	16
Percent Other Types of Communities	66	67	72	72	66	64
Percent in Private Schools	13	16	11	12	13	25
Percent in Public Schools	87	84	89	88	87	75
Correlation Between Proficiency and Student Home Background Index	.36	.36	.49	.50	.34	.36

Percentages may not total 100 percent due to rounding error. Jackknife standard errors were not computed for this basic overview.

Essentially, the characteristics of the samples were comparable, with the exception of the percentages of students identifying themselves as White or Hispanic at grades 4 and 8 and the percentages of grade 12 students in private schools or in advantaged urban communities. These few discrepancies between the original weighted sample and the unweighted subsample were a result of the weighting procedures, rather than the effects of subsampling. The original NAEP sample was weighted to account for oversampling students from schools with relatively high concentrations of Black and/or Hispanic students and students from private schools. These populations were oversampled to increase precision in estimating the assessment results.

The importance of using weights in analyses, such as the hierarchical linear modeling used in this report, depends on the strength of the relationship between the values of the weights and the values of the variables of interest. If there is no relationship, so that the distribution of the variable of interest does not depend on the value of the weight, then the unweighted analysis will produce unbiased results and will have more precision than the weighted results (see Cochran, 1977, Section 5A.2; Kish, 1965, Section 11.7).[58] The statistics in Table A.2, in particular, the average proficiency and the correlation between proficiency and student home background index, suggest that the effects of the weights might be minimal. To determine the probable effect of weighting on the analyses in this report, the hierarchical analyses summarized in Table 2.3, which consider predictors of school effectiveness in mathematics for grade 12 students, were rerun using the subsample data but using the school-level weights to reduce the effect of oversampling on the analysis. The values of proportion of variance explained (R^2) were identical to two decimal places between the weighted analysis and the unweighted analysis in Table 2.3. Furthermore, the regression coefficient estimates were comparable, both in terms of their magnitude and their significance levels. Consequently, the use of weights and subsampling appears to have a minimal effect on the hierarchical analyses shown in this report.

The teacher questionnaire data were not used in the analyses because there are typically few teachers of mathematics at each grade within a school. For example, small schools may even only have one mathematics teacher across several grades (i.e., grades 4 through 6). The limited number of teachers per school complicates the task of performing informative multilevel analyses involving teachers.

[58] Cochran, W. F., *Sampling Techniques* (New York, NY: John Wiley & Sons, 1977).

Kish, L., *Survey Sampling* (New York, NY: John Wiley & Sons, 1965).

The Hierarchical Models. To identify the most and least effective schools for the outlier analysis, a two-level model was used that controlled for students' home background at the first level and school socioeconomic indicators at level 2. The levels are shown below, although the HLM program[59] employed a single-equation regression model where the terms for the level 2 model were substituted into the level 1 equation.

Level 1 — Students: $y_{ij} = \beta_{0j} + \beta_{1j}HB_{ij} + e_{ij}$,

y_{ij} is mathematics proficiency for student i in school j,

β_{0j} is the intercept for school j,

β_{1j} is the regression coefficient (or slope) relating student home background to proficiency, $= \beta_1$ for $j = 1, ..., n$

HB_{ij} is home background for student i in school j (based on a composite of reading materials in the home parents' education level, and number of parents in the home),

e_{ij} is a random error assumed normally distributed with a variance σ that is constant across individuals and schools.

Level 2 — Schools: $\beta_{0j} = \gamma_{00} + \gamma_{01}SES_j + U_j$,

where,

γ_{00} and γ_{01} are school-level regression coefficients,

SES_j is the school SES index for school j (based on a composite of student home background aggregated to the school level, percentages of students in subsidized lunch program, and size and type of community)

U_j is the deviation of school j from expected effectiveness given the school's level of SES.

(The intercept β_{0j} was modelled by school-level variables and a random component. In contrast, the HB slope, β_{1j} was modelled by a constant term and no error term, indicating that it is held fixed

[59] Bryk, A., Raudenbush, S., Seltzer, M. and Congdon, R., *An Introduction to HLM: Computer Program and User's Guide* (University of Chicago, Department of Education, 1989).

across schools. A maximum likelihood test of model fit indicated that the simpler model with a fixed slope fitted the data as well as a model with a random slope.)

β_{0j} represents the mean mathematics proficiency of a school adjusted for the effects of student home background. The term of most interest is the school-level residual, U_j, which represents a school effect controlled for students' home background and school-level SES indicators. From the second equation, the residual can be algebraically solved for,

$$U_j = \beta_{0j} - (\gamma_{00} + \gamma_{01}SES_j),$$

demonstrating that the residual represents the adjusted school mean with school SES effects subtracted out. The residual effects, then, are departures from the expected school performance. A school with a large positive residual had a higher mean proficiency than would be expected, on the average, for schools with students from similar backgrounds and that have similar SES characteristics. A school with a large negative residual had a lower mean than would be expected.

The most effective schools were defined as those with residuals at least one standard deviation *above* the mean and least effective schools were defined as those with residuals one or more standard deviations *below* the mean. For all three grades, of the two groups each comprised about 15 percent of the total number of schools.

The blockwise analyses of the relative effects of various school factors related to school effectiveness also used a two-level model, with level 1 being identical to that used for identifying outlier schools. That is, level 1 represented each school's mean mathematics proficiency adjusted for the effect of each student's home background. However, for this analysis, level 2 is represented as follows:

Level 2 — Schools:

$\beta_{0j} = \gamma_{00} + \gamma_{01}W_{1j} + \ldots + \gamma_{0i}W_{ij} + U_{0j}$
$\beta_{1j} = \gamma_{10},$

for school j where,

β_{0j} is the adjusted proficiency of school j,

β_{1j} is the proficiency on home background regression coefficient from level 1, which is constant over all j schools,

$\gamma_{00}, \gamma_{01}, ..., \gamma_{0i}$ and γ_{10}	are school-level regression coefficients, and
$W_{1j}, ..., W_{ij}$	are school-level predictors
U_{0j}	is a normally distributed random error with variance τ^2. This variance is constant across schools and is independent of the first-level error term.

The HLM program offers the option of centering predictor variables in the first-level (or within-group) model. Since the first-level model can be thought of as a within-group regression, centering predictor variables has the effect of defining the intercept term as a simple group mean. This option is especially useful if group means are to be modeled at the between-group level. For the analyses presented in this report, the within-group predictor is a variable representing students' home background and it is not centered. By not centering the predictor, the intercept term represents the school mean adjusted for students' home background. The adjusted mean represents school effectiveness, the average proficiency of the school with the effect of student factors external to the school taken out. It is school effectiveness that is modeled at the between-school level. For each of the three grades, the amount of total variance between schools was 20 to 21 percent of the total variance of student proficiency adjusted for home background.

As discussed earlier in this Appendix, to increase accuracy, NAEP constructs sets of plausible values designed to represent the distribution of proficiency in the population assessed. In so doing, five imputed values are produced for each student, with the variance of these imputed values indicating the amount of error due to the imprecision of individual measurement. Although the two-level HLM program has been modified to automatically incorporate imputation variance in all estimates and statistical tests,[60] this software was not available in time for the studies reported in Chapter Two. The two-level hierarchical analyses involving imputed proficiency were run five times, once for each imputed value of proficiency. Imputation error was factored into the relevant estimates and statistical tests as recommended by Mislevy.[61]

---

[60] Arnold, C., "Using Hierarchical Linear Models Using NAEP Data." Paper presented at the 1993 annual meeting of the American Educational Research Association, Atlanta, GA.

[61] Mislevy, R., "Randomization Based Inference About Latent Variables from Complex Samples," *Psychometrika*, 56(2), 177-196, 1991.

The Analysis of the Most and Least Effective Schools. For each grade, after the 15 percent most effective and the 15 percent least effective schools were identified, a large set of background variables was subjected to univariate t-tests to determine if there were significant differences between the two groups of schools. The set of variables tested was chosen on the basis of exploratory analyses of correlations with average school proficiencies and breakdowns of school proficiency by levels defined by the variables. The variables that evidenced significant univariate differences were then used as dependent variables in a multivariate analysis of variance (MANOVA), which tested for differences between the most and least effective schools. The MANOVA demonstrated that the most and least effective groups of schools were significantly different with respect to the set of dependent variables with a global alpha level of .05.

At grade 4, the following formed the set of variables subject to significance testing. Only the first six showed significant differences between the two groups of schools and were included in Table 2.1 (see Chapter 2).

- Hours of television students watched per day
- Students were tested weekly or monthly in mathematics class
- School was in the Southeast region of the country
- Average number of times students changed schools in the last two years
- School was a private versus a public school
- How positive were relations between students and teachers
- Percentage of White, Asian/Pacific Islander, and Hispanic students
- Average number of pages students read per day for schoolwork
- Frequency of doing mathematics problems
- School was in the Northeast region of the country
- Students assigned to mathematics class by ability
- School had homework requirements
- Changes in the student body served as impetus to change curriculum or instruction
- Percentage of students who remained in school all year

- Percentage of students retained in fourth grade
- National reports served as impetus to change curriculum or instruction
- Seriousness of problems in the school (sum of several questions)
- Seriousness of problem with teacher absenteeism
- Seriousness of problem with student absenteeism
- Seriousness of problem with physical conflicts among students
- Seriousness of problem with racial/cultural conflicts
- Seriousness of problem with student health
- How positive were attitudes in the school (sum of several questions)
- How positive was teacher morale
- How positive were students' attitudes towards academic achievement
- How positive were teachers' attitudes towards academic achievement
- How positive were parents' attitudes towards academic achievement
- How positive was regard for school property

At grade 8, the following variables were considered for significance testing. The first 11 showed significant differences between the groups of most effective and least effective schools and were included in Table 2.1.

- Percentage of students who planned to take geometry in ninth grade
- School was in the Southeast region of the country
- Percentage of students currently enrolled in algebra
- Hours of television students watched per day
- Seriousness of problem with physical conflicts among students
- Average number of times students changed schools in the last two years
- Students were tested weekly or monthly in mathematics class

- Percentage of White, Asian/Pacific Islander, and Hispanic students
- How often, on average, students used calculators
- How positive were students' attitudes towards academic achievement
- How positive were parents' attitudes towards academic achievement
- Average number of pages students read per day for schoolwork
- Frequency of doing mathematics problems
- Percentage of students doing worksheets weekly or monthly
- Number of external factors that served as an impetus to change the curriculum or instructional practices
- District tests served as impetus to change curriculum or instruction
- State-mandated testing served as impetus to change curriculum or instruction
- National reports served as impetus to change curriculum or instruction
- Change in budget served as impetus to change curriculum or instruction
- The percentage of students who remained in school all year
- Seriousness of problems in the school (sum of several questions)
- Seriousness of problem with student tardiness
- Seriousness of problem with student absenteeism
- Seriousness of problem with class cutting
- Seriousness of problem with racial/cultural conflicts
- Seriousness of problem with student health
- Percentage of students absent on an average day
- Seriousness of problem with teacher absenteeism
- Percentages of students retained in eighth grade
- How positive were attitudes in the school (sum of several questions)

- How positive was teacher morale
- How positive were teachers' attitudes toward academic achievement
- How positive was regard for school property
- How positive were relations between students and teachers

At grade 12, the following variables formed the larger set of variables considered for significance testing. The first 15 variables showed significant differences between the most effective and least effective schools and were included in Table 2.1.

- Average number of advanced mathematics classes students had taken
- Average number of pages students read per day for schoolwork
- Frequency of doing mathematics problems
- How often, on average, students used calculators
- Hours of television students watched per day
- How positive were students' attitudes toward academic achievement
- Percentage of students who took algebra by ninth grade
- Seriousness of problem with student absenteeism
- How positive was teacher morale
- Percentage of students attending college in last year's graduating class
- How positive were parents' attitudes toward academic achievement
- How positive were relations between students and teachers
- Seriousness of problem with physical conflicts among students
- How positive was regard for school property
- Number of external factors that served as an impetus to change the curriculum or instructional practices
- Students were tested weekly or monthly in mathematics class

- School offers advanced mathematics courses
- School was a private versus a public school
- How positive were attitudes in the school (sum of several questions)
- Seriousness of problems in the school (sum of several questions)
- School had homework requirements
- Seriousness of problem with teacher absenteeism
- Percentage of students who remained in school all year
- Seriousness of problem with student tardiness
- Seriousness of problem with class cutting
- Seriousness of problem with racial/cultural conflicts
- Seriousness of problem with student health
- Percentage of students absent on an average day
- How positive were teachers' attitudes toward academic achievement
- State testing mandates served as impetus to change curriculum or instruction
- Percentage of White, Asian/Pacific Islander, and Hispanic students
- School was in the Southeast region of the country

Testing the Significance of Incremental Increases in the Proportion of Variance Explained in the Blockwise HLM Analyses. Although the order of inclusion of the blocks into the model was determined by nonempirical as well as empirical considerations, it is of some interest to determine if the incremental increase in the proportion of variance explained (R^2) incurred by adding each additional block to the model is statistically significant. This can be determined through the following incremental F testing procedure. Let R_1^2 be the proportion of variance explained by including Block 1 in the regression model and let R_2^2 be the proportion of variance explained by including both Block 1 and Block 2 in the model. Then the statistical significance of the increase in the proportion of variance explained by adding Block 2 into the model, over that obtained only by Block 1, is

$$F_{2|1} = [(n - p_1 - p_2)(R_2^2 - R_1^2)] / [p_2 (1 - R_2^2)]$$

where n is the number of schools, p_1 is the number of predictors in Block 1 and p_2 is the number of predictors in Block 2. $F_{2|1}$ has a F distribution with p_2 and $(n - p_1 - p_2)$ degrees of freedom. (For a discussion of such tests, see, for example, Tabachnick and Fidell, 1989, page 157 and Graybill, 1976, page 446.)[62]

These incremental F-tests were applied to each of the block-wise R^2 statistics in Tables 2.2, 2.3, and 2.4. The statistics were significant in every case but one, the increase in variance due to adding Private/Public School to Mathematics Curriculum/Instruction at grade 4. This had an F statistic significant at the 8 percent level.

Technical Issues in the Application of HLM to NAEP Data

The use of hierarchical linear models (HLM) is an attempt to make NAEP data more interpretable than the univariate statistics and tables that are common to NAEP reports. Hierarchical linear models are used to provide valid descriptions of processes that occur in different natural groupings of individuals in the population, e.g., of students grouped within classrooms, students grouped within schools, and schools grouped within the nation. The HLM model can take many forms. When considering a specific model, informed researchers often disagree about what are the proper underlying assumptions that should be made. Models are created to test out different assumptions and conceptions about the world. Because of this, no one model is the ideal model of the world.

In the interest of open and broad consideration of issues and in light of some of the methodological concerns with the model in this report expressed by some researchers, the following discussion is presented.

Because NAEP does not specifically address socioeconomic status (SES), several surrogate indicators at the student and school levels were used to adjust school-level performance. A concern was that the three variables used in the student home background composite did not account for all of the meaningful background differences between students. If this is true then the effects of other variables would reflect some of the confounding influences of student background. It was felt by the authors

[62] Tabachnick, B. G. & Fidell, L. S., *Using Multivariate Statistics* (New York, NY: Harper Collins Publishers, Inc., 1989).

Graybill, F. A., *Theory and Application of the Linear Model* (North Scituate, Mass: Duxbury Press, 1976).

that enough of the students' background and school SES variation was accounted for to make conditional effects more interpretable than raw effects in the HLM analysis.

Some researchers assert that restricting the analyses to just student and school levels discounts the interaction of the teacher and/or the classroom processes with the student. As pointed out elsewhere in the report, the exclusion of the teacher questionnaire data, which is the source of many NAEP instructional variables, makes interpretation about the relative importance of those variables an issue essentially beyond the scope of these analyses. There were relatively few instructional variables measured on the student and those that had significant effects were included in the analysis. However, there were not enough teachers per school to allow for a teacher-within-school level of analysis. Similarly, many schools (approximately 27 percent at grade 8) did not have two or more classes, and the number of students per class was small (28 percent with fewer than two students). It also should be noted that because fewer than two-thirds of the high school seniors take mathematics, teacher questionnaire data were not collected at grade 12, but only at grades 4 and 8.

A related issue is whether the conditioning of proficiencies based upon background variables produces unbiased results for school-level factors. In the 1992 NAEP assessment, virtually all student, teacher, and school variables were used for conditioning, which resulted in unbiased estimates of most effects of interest for students. This is even more true for school-level effects since variables aggregated to the school level are more reliable than student-level measures.

Another topic of debate concerns the sample size of students in schools required for stable estimates of within-school effects. Some researchers feel that the sample sizes in this study are too small to produce reliable within-school regression estimates, while others assert that since the regression estimates in this study are pooled within schools, sample size is not a problem. As an example for comparison, research done by Raudenbush and others[63] to study curricular emphasis of teachers within schools used a sample of 304 mathematics classes taught by 74 teachers within 16 schools. That yields an average of five teachers per school and four classes per teacher.

[63] Raudenbush, S., Rowan, B., and, Cheong, Y., Higher Order Instructional Goals in Secondary Schools: Class, Teacher, and School Influences, *American Educational Research Journal*, 30 (No. 3 Fall), 523-553, 1993.

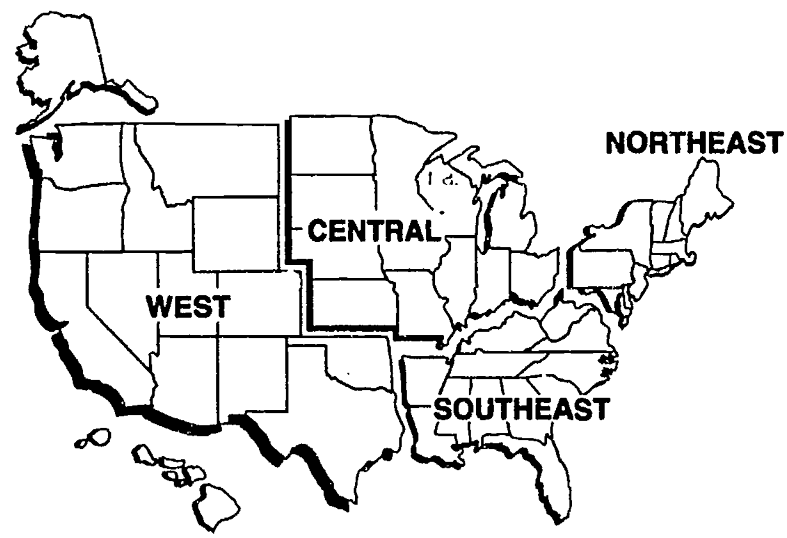
There also are differences of opinion among methodologists regarding the use of outlier analysis techniques to identify effective and ineffective schools. While some researchers consider outliers as anomalous and therefore uninteresting, others focus on these schools as representing contexts where schools are particularly effective or ineffective. In this study, schools were selected that were very high (or low) in effectiveness in that they had very high (or low) average achievement when adjusted for school and home background socioeconomic factors. The study then sought to identify factors associated with extremes in effectiveness. As in the top third/bottom third analysis, the classification of a particular school into a top or bottom group is somewhat arbitrary; for example, some schools in the top group are going to be similar to some schools in the middle group. Some researchers have contended that in the absence of a sharp boundary between the middle group and the extreme groups, the middle group should be included in the analysis. An example of an outlier analysis can be found in a school desegregation report for the Department of Health, Education and Welfare[64], in which the 14 percent top schools on a race relations measure were chosen as were the bottom 13 percent on the measure. These outlier schools were then analyzed to determine which school factors were related to achievement outcomes.

NAEP Demographic Groups

This report contains results for the nation, participating states, and groups of students within the nation defined by shared characteristics. The following are the definitions for subgroups as defined by region, race/ethnicity, size and type of community, and type of school.

Region. The United States has been divided into four regions: Northeast, Southeast, Central, and West. States in each region are shown on the following map.

[64] Forehand, G., Ragosta, M., and, Rock, D., School Conditions and Outcomes in High Schools, Chapter 4 in: Conditions and Processes of Effective School Desegregation, Final Report to the U.S. Department of Health, Education, and Welfare, Educational Testing Service, PR-76-23, 105-140, 1976.



Race/Ethnicity. Results are presented for students of different racial/ ethnic groups based on the students' self-identification of race/ethnicity according to the following mutually exclusive categories: White, Black, Hispanic, Asian/Pacific Islander, and American Indian (including Alaskan Native). Based on statistically determined criteria, at least 62 students in a particular subpopulation must participate in order for the results for that subpopulation to be considered reliable. However, the data for all students, regardless of whether their racial/ethnic group was reported separately, were included in computing the overall national or state level results.

Type of Community. Results are provided for four mutually exclusive community types — advantaged urban, disadvantaged urban, extreme rural, and other — as described below. According to information about parents' occupation obtained from the Principal's Questionnaire completed by each sampled school, indices are developed such that for each assessment approximately the 10 percent of the most extreme advantaged urban, disadvantaged urban, and rural schools are classified into the first three categories. The remaining approximately 70 percent of the schools are classified into the "other" category.

Advantaged Urban: Students in this group reside in metropolitan statistical areas and attend schools where a high proportion of the students' parents are in professional or managerial positions.

Disadvantaged Urban: Students in this group reside in metropolitan statistical areas and attend schools where a high proportion of the students' parents are on welfare or are not regularly employed.

Extreme Rural: Students in this group do not reside in metropolitan statistical areas. They attend schools in areas with a population below 10,000 where many of the students' parents are farmers or farm workers.

Other: Students in the "Other" category attend schools in areas other than those defined as advantaged urban, disadvantaged urban, or extreme rural.

Type of School. For the nation, results are presented separately for public-school students and for private-school students, both those attending Catholic schools and other types of private schools combined.

Estimating Variability

Because the statistics presented in this report are estimates of group and subgroup performance based on samples of students, rather than the values that could be calculated if every student in the nation answered every question, it is important to have measures of the degree of uncertainty of the estimates. Two components of uncertainty are accounted for in the variability of statistics based on proficiency: the uncertainty due to sampling only a relatively small number of students and the uncertainty due to sampling only a relatively small number of mathematics questions. The variability of estimates of percentages of students having certain background characteristics or answering a certain cognitive question correctly is accounted for by the first component alone.

In addition to providing estimates of percentages of students and their proficiency, this report also provides information about the uncertainty of each statistic. Because NAEP uses complex sampling procedures, conventional formulas for estimating sampling variability that assume simple random sampling are inappropriate and NAEP uses a jackknife replication procedure to estimate standard errors. The jackknife standard error provides a reasonable measure of uncertainty for any information about students that can be observed without error, but each student typically responds to so few items within any content area that the proficiency measurement for any single student would be imprecise. In this case, using plausible values technology makes it possible to describe the performance of groups and subgroups of students, but the underlying imprecision that makes this step necessary adds an additional component of variability to statistics based on NAEP proficiencies.[65]

The reader is reminded that the standard error estimates provided with the statistics in this report appropriately take into account uncertainty due to sampling and due to imprecision of individual measurement. NAEP results, like those from all surveys, are also subject to other kinds of errors including the effects of necessarily imperfect adjustment for student and school nonresponse and other largely unknowable effects associated with the particular instrumentation and data collection methods used. Nonsampling errors can be attributed to a number of sources: inability to obtain complete information about all selected students in all selected schools in the sample (some students or schools refused to participate, or students participated but answered only certain items); ambiguous definitions; differences in interpreting questions; inability or unwillingness to give correct information; mistakes in recording, coding, or scoring data; and other errors of collecting, processing, sampling, and estimating missing data. The extent of nonsampling errors is difficult to estimate. By their nature, the impacts of such error cannot be reflected in the data-based estimates of uncertainty provided in NAEP reports.

Drawing Inferences from the Results

The use of *confidence intervals*, based on the standard errors, provides a way to make inferences about the population means and proportions in a manner that reflects the uncertainty associated with the sample estimates. An estimated sample mean proficiency $\pm$ 2 standard errors represents a 95 percent confidence interval for the corresponding population quantity. This means that with approximately 95 percent certainty, the average performance of the entire population of interest is within $\pm$ 2 standard errors of the sample mean.

As an example, suppose that the average mathematics proficiency of students in a particular group was 256, with a standard error of 1.2.

[65] For further details, see Eugene G. Johnson, "Considerations and Techniques for the Analysis of NAEP Data" in *Journal of Educational Statistics* (December 1989).

A 95 percent confidence interval for the population quantity would be as follows:

$$\text{Mean} \pm 2 \text{ standard errors} = 256 \pm 2 \cdot (1.2) = 256 \pm 2.4 = 256 - 2.4 \text{ and } 256 + 2.4 = 253.6, 258.4$$

Thus, one can conclude with 95 percent certainty that the average proficiency for the entire population of students in that group is between 253.6 and 258.4.

Similar confidence intervals can be constructed for percentages, provided that the percentages are not extremely large (greater than 90) or extremely small (less than 10). For extreme percentages, confidence intervals constructed in the above manner may not be appropriate. However, procedures for obtaining accurate confidence intervals are quite complicated. Thus, comparisons involving extreme percentages should be interpreted with this in mind.

To determine whether there is a real difference between the mean proficiency (or proportion of a certain attribute) for two groups in the population, one needs to obtain an estimate of the degree of uncertainty associated with the difference between the proficiency means or proportions of these groups for the sample. This estimate of the degree of uncertainty — called the standard error of the difference between the groups — is obtained by taking the square of each group's standard error, summing these squared standard errors, and then taking the square root of this sum.

Similar to the manner in which the standard error for an individual group mean or proportion is used, the standard error of the difference can be used to help determine whether differences exist between groups in the population. The difference between the mean proficiency or proportion of the two groups ±2 standard errors of the difference represents an approximate 95 percent confidence interval. If the resulting interval includes zero, there is insufficient evidence to claim a real difference between groups in the population. If the interval does not contain zero, the difference between groups is statistically significant (different) at the .05 level.

The procedures described in this section, and the certainty ascribed to intervals (e.g., a 95 percent confidence interval), are based on statistical theory that assumes that only one confidence interval or test of statistical significance is being performed. When one considers sets of confidence intervals, like those for the average proficiency of all participating states and territories, statistical theory indicates that the certainty associated with the entire set of intervals is less than that attributable to each individual comparison from the set. If one wants to hold the certainty level for a specific set of comparisons at a particular level (e.g., .95), adjustments (called multiple-comparisons procedures) need to be made.

The standard errors for means and proportions reported by NAEP are statistics and subject to a certain degree of uncertainty. In certain cases, typically when the standard error is based on a small number of students or when the group of students is enrolled in a small number of schools, the amount of uncertainty associated with the standard errors may be quite large. Throughout this report, estimates of standard errors subject to a large degree of uncertainty are designated by the symbol "!". In such cases, the standard errors — and any confidence intervals or significance tests involving these standard errors — should be interpreted cautiously.

ACKNOWLEDGMENTS

This report is a culmination of effort by hundreds of individuals who contributed their considerable expertise, energy, and creativity to NAEP's mathematics assessments, from developing and conducting the assessments to analyzing and reporting the results. Most importantly, NAEP is grateful to the students and school staff who made the assessments possible.

The 1992 mathematics assessment was truly a collaborative effort among staff from State Education Agencies, the National Center for Education Statistics (NCES), Educational Testing Service (ETS), Westat, and National Computer Systems (NCS). The program benefited from the contributions of hundreds of individuals at the state and local levels — Governors, Chief State School Officers, State and District Test Directors, State Coordinators, and district administrators — who tirelessly provided their wisdom, experience, and hard work.

The assessment was funded through NCES, in the Office of Educational Research and Improvement of the U.S. Department of Education. Emerson Elliott, Commissioner, provided consistent support and guidance. The staff — particularly Gary Phillips, Sharif Shakrani, Peggy Carr, Alex Sedlacek, Steven Gorman, Eugene Owen, Maureen Treacy, Andrew Kolstad, and Susan Ahmed — worked closely and collegially with ETS, Westat, and NCS staff and played a crucial role in all aspects of the program. The members of the National Assessment Governing Board (NAGB) and the NAGB staff provided advice and guidance throughout.

The Chief State School Officers managed the National Assessment Planning Project that resulted in the mathematics framework and objectives for the assessment. NAEP owes a debt of gratitude to the numerous panelists and consultants who worked so diligently on developing the assessment and providing a frame for interpreting the results, including those who helped create the objectives and develop the assessment instruments.

The NAEP project at ETS resides in the Center for the Assessment of Educational Progress (CAEP) managed by Archie Lapointe. Under the NAEP contract to ETS, Ina Mullis served as the Project Director. Stephen Koffler managed test development and state services. The mathematics test development team led by Chancey Jones and Jeffrey Haberstroh, and including Gail Chapman, Jane Kupin, Patty Klag, Jean MacFarland, Jane Maroney, Pat Meier, Melissa White, and Beverly Whittington, deserves special credit for its dedication. NAEP also owes a debt of gratitude to members of the Mathematics Item Development Panel who worked so

diligently on developing the assessment and the Background Questionnaire Panel for their work in developing the framework and questionnaires.

Jules Goodison managed the operational aspects with the strong support of John Olson, who also provided invaluable help in managing state services and reporting. Sampling and data collection activities were carried out by Westat under the direction of Renee Slobasky, Nancy Caldwell, and Keith Rust. Printing, distribution, scoring, and processing activities were conducted by NCS, under the supervision of John O'Neill, Judy Moyer, Diane Smrdel, Lavonne Mohn, Brad Thayer, and Mathilde Kennel.

Mathematics statistical and psychometric procedures and implementation were led by Frank Jenkins under the direction of Eugene Johnson. Frank Jenkins managed procedures for the national mathematics assessment, including the hierarchical analyses. Major contributions were made by James Carlson, Hua Hua Chang, Angela Grima, Jo-lin Liang, Eiji Muraki, Spencer Swinton, and Ming-mei Wang. John Barone managed the data analysis and computerized reporting systems. David Freund and Edward Kulick performed the mathematics data analysis, assisted by Drew Bowker, Bruce Kaplan, Steve Isham, and Yim Fai Fong. Ina Mullis, Frank Jenkins, and Eugene Johnson conceptualized and wrote the text for the report. Many thanks are provided to the numerous reviewers, internal and external, who suggested improvements to successive drafts, especially Kent Ashworth and Jim Carlson. Carol Carlson edited the report and Sharon Davis-Johnson provided the excellent desktop publishing skills essential to the project. Kent Ashworth, who also manages information services, coordinated the cover design and production of the report. The ETS Publications Division provided expert composition.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1994 - 300-807 - 814/23120

ISBN 0-16-045331-3

90000

9 780160 453311

For sale by the U.S. Government Printing Office
Superintendent of Documents, Mail Stop: SSOP, Washington, DC 20402-9328
ISBN 0-16-045331-3

DEPARTMENT OF EDUCATION

UNITED STATES OF AMERICA

NCES 94-701