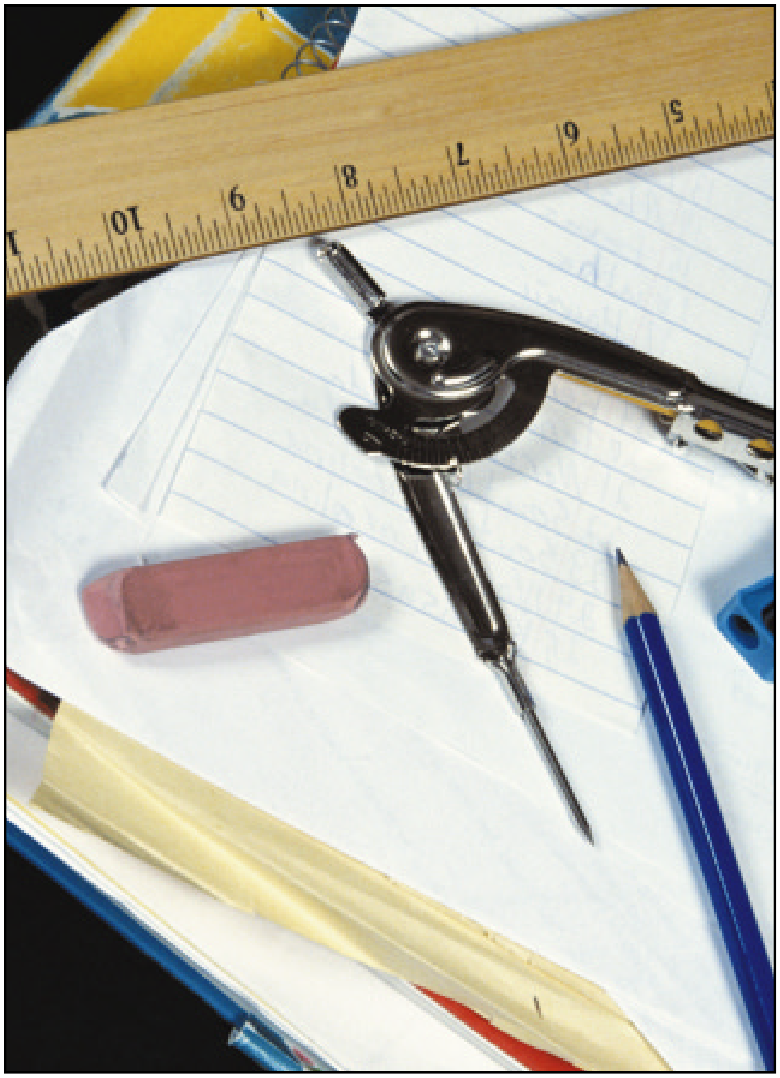
N A T I O N A L   C E N T E R   F O R   E D U C A T I O N   S T A T I S T I C S

NAEP 1996 MATHEMATICS

State Report for New Jersey



THE NATION'S REPORT CARD naep

U.S. DEPARTMENT OF EDUCATION OFFICE OF EDUCATIONAL RESEARCH AND IMPROVEMENT

Findings from the National Assessment of Educational Progress

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 established the National Assessment Governing Board (NAGB) to formulate policy guidelines for NAEP. The Board is responsible for selecting the subject areas to be assessed from among those included in the National Education Goals; for setting appropriate student performance levels; for developing assessment objectives and test specifications through a national consensus approach; for designing the assessment methodology; for developing guidelines for reporting and disseminating NAEP results; for developing standards and procedures for interstate, regional, and national comparisons; for determining the appropriateness of test items and ensuring they are free from bias; and for taking actions to improve the form and use of the National Assessment.

The National Assessment Governing Board

NATIONAL CENTER FOR EDUCATION STATISTICS

NAEP 1996 MATHEMATICS STATE REPORT

for

NEW JERSEY

Clyde M. Reese
Laura Jerry
Nada Ballator

In collaboration with
Peggy Carr, Jeff Haberstroh
Paul Koehler, Phillip Leung
Mary Lindquist, and John Mazzeo

June 1997

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

Prepared by Educational Testing Service under a cooperative agreement with the National Center for Education Statistics.

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

Office of Educational Research and Improvement
Ramon C. Cortines
Acting Assistant Secretary

National Center for Education Statistics
Pascal D. Forgione, Jr.
Commissioner

Education Assessment Group
Gary W. Phillips
Associate Commissioner

June 1997

SUGGESTED CITATION
Reese, C.M., Jerry, L., and Ballator, N.
NAEP 1996 Mathematics State Report for New Jersey,
Washington, DC: National Center for Education Statistics, 1997.

FOR MORE INFORMATION
Contact:
Arnold A. Goldstein
202-219-1741

For ordering information on this report, write:

National Library of Education
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, DC, metropolitan area call 202-219-1651).

This report also is available on the World Wide Web: http://www.ed.gov/NCES/naep

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.

Table of Contents

HIGHLIGHTS

Monitoring the performance of students in subjects such as mathematics is a key concern of the citizens, policy makers, and educators concerned with educational reform efforts. The 1996 National Assessment of Educational Progress (NAEP) in mathematics (as well as the two previous NAEP assessments in mathematics in 1990 and 1992) assessed the current level of mathematical achievement as a mechanism for informing education reform. This report contains results for public school students only for those years in which New Jersey participated and for which minimum participation rate guidelines were met.

What Is NAEP?

The National Assessment of Educational Progress (NAEP) is the only nationally representative and continuing assessment of what students in the United States know and can do in various academic subjects. NAEP is authorized by Congress and directed by the National Center for Education Statistics of the U.S. Department of Education. The National Assessment Governing Board (NAGB), an independent body, provides policy guidance for NAEP.

Since its inception in 1969, NAEP's mission has been to collect, analyze, and produce valid and reliable information about the academic performance of students in the United States in various learning areas. In 1990, the mission of NAEP was expanded to provide state-by-state results on academic achievement. Participation in the state-by-state NAEP is voluntary and has grown from 40 states and territories in 1990 to 48 in 1996.

NAEP has also become a valuable tool in tracking progress towards the National Education Goals. The subjects assessed by NAEP are those highlighted at the 1989 Education Summit and later legislation.[1] The NAEP 1996 assessment in mathematics marks the third time the subject has been assessed with the new framework in the 1990s, enabling policy makers and educators to track mathematics achievement since the release of the National Council of Teachers of Mathematics (NCTM) *Curriculum and Evaluation Standards for School Mathematics*[2] in 1989.

[1] Executive Office of the President. *National Goals for Education*. (Washington, DC: Government Printing Office, 1990); Goals 2000: Educate America Act, Pub. L. No. 103-227 (1994).

[2] National Council of Teachers of Mathematics. *Curriculum and Evaluation Standards for School Mathematics*. (Reston, VA: NCTM, 1989).

NAEP 1996 Mathematics Assessment

The NAEP mathematics assessment has been in constant evolution since its inception in 1973. Major changes took place in the 1990s to complement the *Curriculum and Evaluation Standards for School Mathematics*, that was published by the National Council of Teachers of Mathematics (NCTM) in 1989. The NAEP 1990 mathematics assessment saw the inclusion of short constructed-response questions — questions that asked students to provide the answer they calculated for a numerical problem or to write a sentence or two describing the solution to a problem. Also added in 1990 were a number of questions on which students could use calculators, protractors, or rulers.

In 1992 the assessment included an increased number of short constructed-response questions and, for the first time, contained extended constructed-response questions. Extended constructed-response questions required students to produce both a solution and a short paragraph describing the solution or its interpretation in the context of the task. As such, these questions served as indicators of students' growth in the areas of reasoning, communication, and problem solving — important processes receiving heavy emphasis in the NCTM *Standards*.

In 1996 the NAEP mathematics assessment continued to be revised, most notably by continuing to increase the use of constructed-response questions. In 1990, students spent about 30 percent of testing time on constructed-response questions. By 1992, this percentage had increased to 35 percent, and in 1996 it exceeded 50 percent of the time spent by students on the assessment.

The 1996 assessment maintained the same five content strands used for the 1990 and 1992 assessments — Number Sense, Properties, and Operations; Measurement; Geometry and Spatial Sense; Data Analysis, Statistics, and Probability; and Algebra and Functions. Two of these strands, Number Sense, Properties, and Operations and Geometry and Spatial Sense, were revised to reflect the NCTM *Standards*' emphases on developing and assessing students' abilities to make sense of both number/operation and spatial settings.

The changes made to the NAEP 1996 mathematics assessment refined and sharpened the assessment to reflect more adequately recent curricular emphases and objectives; to include what teachers, mathematicians, and measurement experts think should be in the assessment; and to maintain the connection with the 1990 and 1992 assessments to permit the measurement of trends in student performance since 1990.

Tables H.1 and H.2 show the distribution of mathematics scores and the percentage of students at or above the *Basic*, *Proficient*, and *Advanced* achievement levels for fourth-grade students attending public schools in New Jersey in 1996.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE H.1 — GRADE 4

Distribution of Mathematics Scale Scores for Public School Students

	Average Scale Score	10th Percentile	25th Percentile	50th Percentile	75th Percentile	90th Percentile
Grade 4						
New Jersey	227 (1.5)	186 (2.8)	207 (1.3)	229 (1.7)	249 (1.6)	266 (1.4)
Northeast	226 (2.6)	181 (4.4)	206 (5.0)	229 (2.4)	250 (1.6)	266 (2.8)
Nation	222 (1.0)	180 (1.5)	201 (1.3)	224 (1.3)	244 (1.0)	261 (1.0)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details).
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE H.2 — GRADE 4

Percentage of Public School Students Attaining Mathematics Achievement Levels

	Advanced	At or Above Proficient	At or Above Basic	Below Basic
Grade 4				
New Jersey	3 (0.7)	25 (1.7)	68 (2.1)	32 (2.1)
Northeast	3 (1.0)	26 (1.6)	68 (3.5)	32 (3.5)
Nation	2 (0.3)	20 (1.0)	62 (1.4)	38 (1.4)

The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details).
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

Major Findings for New Jersey

- The average mathematics scale score for fourth graders in New Jersey was 227.[3] This average was higher than that for the nation (222).
- In terms of achievement levels established for the NAEP mathematics assessment, 25 percent of the fourth-grade students in New Jersey performed at or above the *Proficient* level.[4] This percentage was larger than that of students nationwide (20 percent).
- From 1992 to 1996 in grade 4, the average scale score of students in New Jersey did not change significantly while that of students across the nation increased.

[3] The NAEP mathematics scale ranges from 0 to 500.

[4] The *Proficient* achievement level represents solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.

Major Findings for Student Subpopulations

The preceding section provided a global view of the mathematics performance of fourth-grade students in New Jersey. It is also important to examine the average performance of subgroups within this population. Typically, NAEP presents results for demographic subgroups defined by gender, race/ethnicity, parental education, and location of the school. In addition, in 1996 NAEP collected information on student participation in Title I programs and eligibility for the free/reduced-price lunch component of the National School Lunch Program (NSLP).

The reader is cautioned against using NAEP results to make simple or causal inferences related to subgroup membership. Differences among groups of students are almost certainly associated with a broad range of socioeconomic and educational factors not discussed in NAEP reports and possibly not addressed by the NAEP assessment program.

Results for 1996 related to gender and race/ethnicity are highlighted below. More complete results for the various demographic subgroups examined by the NAEP mathematics assessment can be found in Chapters 2 and 4 of this report, the *NAEP 1996 Mathematics State Report for New Jersey*.

- The average mathematics scale score of fourth-grade males was higher than that of females in New Jersey; nationwide, however, the performance of males did not differ significantly from that of females.

- At the fourth grade, White students in New Jersey had an average mathematics scale score that was higher than that of Black and Hispanic students but was lower than that of Asian/Pacific Islander students.

Finding a Context for Understanding Students' Mathematics Performance in Public Schools

The mathematics performance of students in New Jersey may be better understood when viewed in the context of the environment in which the students are learning. This educational environment is largely determined by school characteristics, by characteristics of mathematics instruction in the school, by home support for academics and other home influences, and by the students' own views about mathematics. Information about this environment is gathered by means of the questionnaires completed by principals and teachers as well as questions answered by students as part of the assessment.

Because NAEP is administered to a sample of students that is representative of the fourth- and eighth-grade student populations in the schools of New Jersey, NAEP results provide a view of the educational practices in New Jersey which may be useful for improving instruction and setting policy. However, despite the richness of context provided by the NAEP results, it is very important to note that NAEP data cannot establish a cause-and-effect relationship between educational environment and student scores on the NAEP mathematics assessment.

The following results are for public school students.

School Characteristics Related to Student Performance[5]

- The percentage of fourth-grade students in New Jersey attending public schools that reported that mathematics was a priority (81 percent) was not significantly different from the national percentage (76 percent).

- The percentage of fourth graders attending public schools in New Jersey that reported that absenteeism was a moderate to serious problem (9 percent) was not significantly different from that of fourth graders across the nation (13 percent). The percentage of students in New Jersey public schools reporting that absenteeism was a moderate to serious problem did not change significantly from 1992 (8 percent) to 1996 (9 percent).

Classroom Practices[6]

- A small percentage of the fourth-grade students in New Jersey (5 percent) had mathematics teachers who reported being very knowledgeable about the NCTM *Standards*. This percentage was smaller than the percentage whose teachers reported having little or no knowledge of the *Standards* (42 percent).

---

[5] More detailed results related to school characteristics can be found in Chapter 5 of this report, the *NAEP 1996 Mathematics State Report for New Jersey*.

[6] More detailed results related to classroom practices can be found in Chapter 6 of this report, the *NAEP 1996 Mathematics State Report for New Jersey*.

- The percentage of fourth graders in New Jersey whose teachers reported spending four hours a week or more on mathematics instruction (71 percent) was not significantly different from the percentage for the nation (69 percent). The percentage for New Jersey in 1996 did not differ significantly from the percentage in 1992 (69 percent).

- Teachers of 59 percent of the fourth-grade students reported that they addressed the development of reasoning and analytical ability a lot. In contrast, 5 percent had teachers who reported spending little or no time addressing this topic.

- According to their teachers, 6 percent of the fourth graders in New Jersey were asked to write about solving a mathematics problem and 42 percent were asked to discuss solutions with other students almost every day. By comparison, 23 percent were asked to write about and 4 percent were asked to discuss mathematics solutions never or hardly ever.

- According to their teachers, 0 percent of the fourth graders in New Jersey were not assigned any mathematics homework each day. In addition, almost all of the students were assigned 15 minutes (38 percent) or 30 minutes (56 percent) of homework each day.

- Less than half of the fourth graders in New Jersey reported that there was no computer at home (39 percent) and another 28 percent reported never or hardly ever using their home computer to do homework. Less than one fifth of the students reported using a computer at home for homework almost every day (6 percent) or once or twice a week (10 percent).

- Less than half of the fourth graders in New Jersey had teachers who reported that students used a calculator in mathematics class almost every day (6 percent) or once or twice a week (39 percent). Less than one fifth of the students never or hardly ever used a calculator (15 percent).

Influences Beyond School That Facilitate Learning Mathematics[7]

- More than half of the fourth graders (58 percent) said they discussed their schoolwork at home almost every day. This percentage was larger than the percentage who said they never or hardly ever had such discussions (13 percent).

- The percentage of fourth graders in New Jersey who reported watching six or more hours of television a day (22 percent) was not significantly different from the percentage for the nation (20 percent).

- Overall, almost all of the fourth-grade students attended schools where principals characterized parental support as very positive (55 percent) or somewhat positive (41 percent).

[7] More detailed results related to influences beyond the school can be found in Chapter 7 of this report, the *NAEP 1996 Mathematics State Report for New Jersey*.

INTRODUCTION

Improving education is often seen as an important first step as the United States maps out a strategy to remain competitive in an ever-increasing global economy. Mathematics and science education continued to receive considerable attention at the 1996 Governor's Summit in Palisades, New Jersey, where the President and the governors reaffirmed the need to strengthen our schools and to strive for world-class standards.

Monitoring the performance of students in subjects such as mathematics is a key concern of the state and national policy makers and educators who direct educational reform efforts. The 1996 National Assessment of Educational Progress (NAEP) in mathematics is a key source of information on what the nation's students can do and how mathematics achievement has progressed during the 1990s.

What Was Assessed?

The NAEP assessment measures a mathematics domain containing five mathematics strands (number sense, properties, and operations; measurement; geometry and spatial sense; data analysis, statistics, and probability; and algebra and functions). Questions involving content from one or more of the strands are also categorized according to the domains of mathematical abilities and mathematical power. The first of these, mathematical abilities, describes the nature of the knowledge or processes involved in successfully handling the task presented by the question. It may reflect conceptual understanding, procedural knowledge, or a combination of both in problem solving. The second domain, mathematical power, reflects processes stressed as major goals of the mathematical curriculum. Mathematical power refers to the students' ability to reason, to communicate, and to make connections of concepts and skills across mathematical strands, or from mathematics to other curricular areas.

The mathematics framework for the NAEP 1996 assessment is a revision of that used in the 1990 and 1992 assessments. Changes were made to the earlier framework in light of the NCTM Standards and changes taking place in school mathematics programs. The previous NAEP mathematics framework was refined and sharpened so that the 1996 assessment would: (1) more adequately reflect recent curricular emphases and objects and yet (2) maintain a connection with the 1990 and 1992 assessments to measure trends in student performance. Prior to the 1996 assessment, investigations were conducted to ensure that results from the assessment could be reported on the existing NAEP mathematics scale. The conclusion drawn from these investigations was that results from the 1990, 1992, and 1996 assessments could be reported on a common scale and trends in mathematics performance since 1990 examined. Appendix B briefly highlights selected changes in the current NAEP mathematics framework.

The conception of mathematical power as reasoning, connections, and communication has played an increasingly important role in measuring student achievement. In 1990, the NAEP assessment included short constructed-response questions as a way to begin addressing mathematical communication. In 1992, the extended constructed-response questions included on the assessment required students not only to communicate their ideas but also to demonstrate the reasoning they used to solve problems. The 1996 assessment continued to emphasize mathematical power by including constructed-response questions focusing on reasoning and communication and by requiring students to connect their learning across mathematical content strands. These connections were addressed within individual questions reaching across content strands and by families of questions contained within a single content strand.

In real life, few mathematical situations can be clearly classified as belonging to one content strand or another, and few situations require only one fact of mathematics thinking. Therefore, many of the questions are classified in a number of ways. In addition to being classified by all applicable content strands, each question was classified by its assessment of applicable mathematical abilities (procedural knowledge, conceptual understanding, and problem solving) and mathematical powers (reasoning, communication, and connections). The content strands, mathematical abilities, and mathematical power combine to form the framework for the NAEP assessment. (A brief description of the five content strands is presented in Appendix B.)

The framework continued the shift from multiple-choice questions to questions that required students to construct responses. In 1996, more than 50 percent of student assessment time was devoted to constructed-response questions. Two types of constructed-response questions were included — (1) short constructed-response questions that required students to provide answers to computation problems or to describe solutions in one or two sentences, and (2) extended constructed-response questions that required students to provide longer responses when answering the questions.

Who Was Assessed?

Fourth-Grade School and Student Characteristics

Table I.1 provides a profile of the demographic characteristics of the fourth-grade students in New Jersey, the Northeast region, and the nation. This profile is based on data collected from the students and schools participating in the 1992 and 1996 state and national mathematics assessments at grade 4. This report contains results for public school students only for those years in which New Jersey participated and for which minimum participation rate guidelines were met. As described in Appendix A, the state data and the regional and national data are drawn from separate samples.

To ensure comparability across jurisdictions, NCES has established guidelines for school and student participation rates. Appendix A highlights these guidelines, and jurisdictions failing to meet these guidelines are noted in tables and figures in NAEP reports containing state-by-state results. For jurisdictions failing to meet the initial school participation rate of 70 percent, results are not reported.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE I.1 — GRADE 4

Profile of Students in New Jersey, the Northeast Region, and the Nation

Demographic Subgroups		1992	1996
		Public	
		Percentage	
RACE/ETHNICITY			
New Jersey	White	66 (2.2)	59 (2.8)
	Black	14 (1.2)	21 (2.4)
	Hispanic	14 (1.5)	14 (1.6)
	Asian/Pacific Islander	5 (0.8)	5 (0.5)
	American Indian	1 (0.3)	2 (0.3)
Northeast	White	71 (2.9)	67 (2.6)
	Black	17 (2.7)	16 (2.5)
	Hispanic	8 (1.2)	12 (2.1)
	Asian/Pacific Islander	2 (0.7)	2 (0.7)
	American Indian	1 (0.3)	2 (0.4)
Nation	White	69 (0.4)	66 (0.6)<
	Black	17 (0.4)	15 (0.4)
	Hispanic	10 (0.2)	14 (0.4)>
	Asian/Pacific Islander	3 (0.3)	3 (0.2)
	American Indian	2 (0.2)	2 (0.2)
TYPE OF LOCATION *			
New Jersey	Central city	14 (2.7)	15 (4.2)
	Urban fringe/Large town	82 (3.3)	84 (4.4)
	Rural/Small town	3 (1.7)	1 (****)
Nation	Central city	32 (3.0)	28 (2.9)
	Urban fringe/Large town	40 (3.6)	46 (3.6)
	Rural/Small town	28 (2.9)	26 (2.8)
PARENTS' EDUCATION			
New Jersey	Did not finish high school	3 (0.4)	3 (0.5)
	Graduated from high school	11 (0.7)	12 (1.1)
	Some education after high school	8 (0.6)	7 (0.6)
	Graduated from college	47 (1.7)	46 (2.0)
	I don't know	31 (1.3)	32 (1.5)
Northeast	Did not finish high school	4 (0.7)	4 (0.7)
	Graduated from high school	11 (0.9)	12 (1.6)
	Some education after high school	6 (0.6)	8 (0.9)
	Graduated from college	44 (3.2)	40 (2.3)
	I don't know	35 (2.0)	36 (2.1)
Nation	Did not finish high school	4 (0.3)	4 (0.4)
	Graduated from high school	13 (0.6)	13 (0.7)
	Some education after high school	7 (0.4)	7 (0.4)
	Graduated from college	40 (1.1)	38 (1.2)
	I don't know	36 (0.8)	37 (1.0)
GENDER			
New Jersey	Male	51 (1.0)	49 (1.4)
	Female	49 (1.0)	51 (1.4)
Northeast	Male	50 (1.2)	51 (1.4)
	Female	50 (1.2)	49 (1.4)
Nation	Male	50 (0.7)	51 (0.7)
	Female	50 (0.7)	49 (0.7)

(continued on next page)

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE I.1 — GRADE 4 (continued)

Profile of Students in New Jersey, the Northeast Region, and the Nation

Demographic Subgroups		1992	1996
		Public	
		Percentage	
TITLE I			
New Jersey	Participated	--- (--.-)	14 (1.7)
	Did not participate	--- (--.-)	86 (1.7)
Northeast	Participated	--- (--.-)	19 (2.9)
	Did not participate	--- (--.-)	81 (2.9)
Nation	Participated	--- (--.-)	24 (1.5)
	Did not participate	--- (--.-)	76 (1.5)
FREE/REDUCED-PRICE LUNCH			
New Jersey	Eligible	--- (--.-)	33 (2.1)
	Not eligible	--- (--.-)	65 (2.3)
	Information not available	--- (--.-)	2 (1.2)
Northeast	Eligible	--- (--.-)	31 (2.2)
	Not eligible	--- (--.-)	60 (4.6)
	Information not available	--- (--.-)	9 (4.9)
Nation	Eligible	--- (--.-)	34 (1.6)
	Not eligible	--- (--.-)	52 (2.5)
	Information not available	--- (--.-)	13 (3.1)

The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. The percentages for Race/Ethnicity may not add to 100 percent because some students categorized themselves as "Other." * Characteristics of the school sample do not permit reliable regional results for type of location. --- Title I and Free/Reduced-Price Lunch results are not available for the 1992 assessment. **** Standard error estimates cannot be accurately determined.

SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

Schools and Students Assessed

Table I.2 summarizes participation data for schools and students sampled in New Jersey for the 1996 state assessment program in mathematics and previous NAEP state assessments in mathematics.[8]

In New Jersey, 78 public schools participated in the 1996 fourth-grade mathematics assessment. This number includes participating substitute schools that were selected to replace some of the nonparticipating schools from the original sample. The weighted school participation rate after substitution in 1996 was 73 percent for public schools, which means that the fourth-grade students in this sample were directly representative of 73 percent of all the fourth-grade public school students in New Jersey.

In each school, a random sample of students was selected to participate in the assessment. In 1996, on the basis of sample estimates, 2 percent of the fourth-grade public school population were classified as students with limited English proficiency (LEP). At the fourth grade, 9 percent of the students in public schools had an Individualized Education Plan (IEP). An IEP is a plan written for a student who has been determined to be eligible for special education. The IEP typically sets forth goals and objectives for the student and describes a program of activities and/or related services necessary to achieve the goals and objectives.

Schools were permitted to exclude certain students from the assessment, provided that the following criteria were met. To be excluded, a student had to be categorized as LEP or had to have an IEP or equivalent *and* (in either case) be judged incapable of participating in the assessment. The intent was to assess all selected students; therefore, all selected students who were capable of participating in the assessment should have been assessed. However, schools were allowed to exclude those students who, in the judgment of school staff, could not meaningfully participate. The NAEP guidelines for exclusion are intended to assure uniformity of exclusion criteria from school to school. Note that some students classified as LEP and some students having an IEP were deemed eligible to participate and not excluded from the assessment. The students in New Jersey who were excluded from the assessment because they were categorized as LEP or had an IEP represented 6 percent of the public school population in grade 4.

In New Jersey 1,961 public school fourth-grade students were assessed in 1996. The weighted student participation rate was 95 percent for public schools. This means that the sample of fourth-grade students who took part in the assessment was directly representative of 95 percent of the eligible public school student population in participating schools in New Jersey (that is, all students from the population represented by the participating schools, minus those students excluded from the assessment). The overall weighted response rate (school rate times student rate) was 69 percent for public schools. This means that the sample of students who participated in the assessment was directly representative of 69 percent of the eligible fourth-grade public school population in New Jersey.

---

[8] For a detailed discussion of the NCES guidelines for sample participation, see Appendix A of this report or the *Technical Report of the NAEP 1996 State Assessment Program in Mathematics*. (Washington, DC: National Center for Education Statistics, 1997).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE I.2 — GRADE 4

Profile of the Population Assessed in New Jersey

	1992	1996	
	Public	Public	Nonpublic
SCHOOL PARTICIPATION			
Weighted school participation rate before substitution	76%	73%	64%
Weighted school participation rate after substitution	82%	73%	75%
Number of schools originally sampled	119	107	26
Number of schools not eligible	3	1	4
Number of schools in original sample participating	88	78	14
Number of substitute schools provided	22	24	8
Number of substitute schools participating	7	0	2
Total number of participating schools	95	78	16
STUDENT PARTICIPATION			
Weighted student participation rate after makeups	96%	95%	94%
Number of students selected to participate in the assessment	2,532	2,192	353
Number of students withdrawn from the assessment	77	58	0
Percentage of students who were of Limited English Proficiency	4%	2%	0%
Percentage of students excluded from the assessment due to Limited English Proficiency	2%	1%	0%
Percentage of students who had an Individualized Education Plan	8%	9%	3%
Percentage of students excluded from the assessment due to Individualized Education Plan status	3%	5%	0%
Number of students to be assessed	2,322	2,068	353
Number of students assessed	2,231	1,961	334
Overall weighted response rate	79%	69%	70%

In accordance with standard practice in survey research, the results presented in this report were based on calculations that incorporate adjustments for the nonparticipating schools and students. Hence, the final results derived from the sample provide estimates of the mathematics performance for the full population of eligible fourth-grade students in New Jersey. However, in instances where nonparticipation rates are large, these nonparticipation adjustments may not adequately compensate for the missing sample schools and students.

In order to guard against potential nonparticipation bias in published results, the National Center for Education Statistics (NCES) has established minimum participation levels as a condition for the publication of 1996 state assessment program results. NCES also established additional guidelines addressing four ways in which nonparticipation bias could be introduced into a jurisdiction's published results (see Appendix A). In 1996 New Jersey met minimum participation levels for public schools at grade 4. However, New Jersey failed to meet minimum participation levels for public schools at grade 8 and for nonpublic schools at grades 4 and 8. The weighted participation rates for the initial samples of public schools at grade 8 and for nonpublic schools at grades 4 and 8 were less than 70%. Hence, results are included in this report only for public schools at grade 4. New Jersey failed to meet one or more established NCES guidelines for public schools at grade 4. The weighted participation rate for the initial sample was below 85 percent and the weighted school participation rate after substitution was below 90 percent (see Appendix A).

In the analysis of student data and reporting of results, nonresponse weighting adjustments have been made at both the school and student level, with the aim of making the sample of participating students as representative as possible of the entire eligible fourth-grade population. For details of the nonresponse weighting adjustment procedures, see the *Technical Report of the NAEP 1996 State Assessment Program in Mathematics*.

Reporting NAEP Mathematics Results

The NAEP 1996 state assessment program in mathematics provides a wealth of information on the mathematical abilities and skills of the fourth-grade students in participating jurisdictions. To maximize usefulness to policy makers, educators, parents, and other interested parties, the NAEP results are presented both as average scale scores on the NAEP mathematics scale and in terms of the percentage of students attaining NAEP mathematics achievement levels. Thus, NAEP results not only provide information about what students *know and can do,* but also indicate whether their achievement meets expectations of what students *should know and should be able to do*. Furthermore, the descriptions of skills and abilities expected of students at each achievement level help make the reporting of assessment results more meaningful.

The Mathematics Scale

Students' responses to the NAEP 1996 mathematics assessment were analyzed to determine the percentage of students responding correctly to each multiple-choice question and the percentage of students responding in each of several score categories for constructed-response questions. Item response theory (IRT) methods were used to produce across-grade scales that summarized results for each of the five mathematics content strands discussed earlier. Each of the content-strand scales, which range from 0 to 500, was linked to its corresponding scale from 1990 and 1992 through IRT equating.

An overall composite scale was developed by weighting the separate content-strand scales based on the relative importance to each content strand in the NAEP mathematics framework. The resulting scale, which was also linked to the 1990 and 1992 mathematics composite scales, is the reporting metric used in Parts One and Three to present results. (Details of the scaling procedures are presented in Appendix C of this report, in the *NAEP 1996 Technical Report*, and in the *Technical Report of the NAEP 1996 State Assessment Program in Mathematics*.)

Mathematics Achievement Levels

Results for the NAEP 1996 assessment in mathematics are also reported using the mathematics achievement levels that were authorized by the NAEP legislation and adopted by the National Assessment Governing Board. The achievement levels are based on collective judgments about what students *should know and be able to do* relative to the body of content reflected in the NAEP mathematics assessment. Three levels were defined for each grade — *Basic, Proficient,* and *Advanced*. The levels were defined by a broadly representative panel of teachers, education specialists, and members of the general public.

For reporting purposes, the achievement levels for each grade are placed on the NAEP mathematics scale. Figure 1 presents the policy definitions of the achievement levels, while Chapter 3 contains specific descriptions for the levels.

Figure 1. Policy Definitions of NAEP Achievement Levels

Basic	This level denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.
Proficient	This level represents solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.
Advanced	This level signifies superior performance.

It should be noted that setting achievement levels is a relatively new process for NAEP, and it is still in transition. Some evaluations have concluded that the percentage of students at certain levels may be underestimated.[9] On the other hand, critiques of those evaluations have asserted that the weight of the empirical evidence does not support such conclusions.[10] A further review is currently being conducted by the National Academy of Sciences.

The student achievement levels in this report have been developed carefully and responsibly, and the procedures used have been refined and revised as new technologies have become available. Upon review of the available information, the Commissioner of Education Statistics has judged that the achievement levels are in a developmental status. However, the Commissioner and the Governing Board also believe that the achievement levels are useful and valuable for reporting on the educational achievement of students in the United States. Part Two presents results reported in terms of the mathematics achievement levels.

Interpreting NAEP Results

This report describes mathematics performance for fourth graders and compares the results for various groups of students within this population — for example, those who have certain demographic characteristics or who responded to a specific background question in a particular way. The report examines the results for individual demographic groups and for individual background questions. It does not include an analysis of the relationships among combinations of these subpopulations or background questions.

Because the percentages of students in these subpopulations and their average mathematics scale scores are based on samples — rather than on the entire population of fourth graders in a jurisdiction — the numbers reported are necessarily *estimates*. As such, they are subject to a measure of uncertainty, reflected in the *standard error* of the estimate. When the percentages or average scale scores of certain groups are compared, it is essential to take the standard error into account, rather than to rely solely on observed similarities or differences. Therefore, the comparisons discussed in this report are based on *statistical tests* that consider both the magnitude of the difference between the means or percentages and the standard errors of those statistics.

---

[9] General Accounting Office. *Educational Achievement Standards: NAGB's Approach Yields Misleading Interpretations*. (Washington, DC: General Accounting Office, 1993); National Academy of Education. *Setting Performance Standards for Student Achievement*. A Report of the National Academy of Education Panel on the Evaluation of the NAEP Trial State Assessment: An Evaluation of the 1992 Achievement Levels. (Stanford, CA: National Academy of Education, 1993).

[10] Cizek, G. *Reactions to the National Academy of Education report*. (Washington, DC: National Assessment Governing Board, 1993); Kane, M. *Comments on the NAE Evaluation of the NAGB Achievement Levels*. (Washington, DC: National Assessment Governing Board, 1993); *NAEP Reading Revisited: An Evaluation of the 1992 Achievement Levels Descriptions*. (American College Testing, Washington, DC: National Assessment Governing Board, 1993); *Technical Report on Setting Achievement Levels on the 1992 National Assessment of Educational Progress in Mathematics, Reading, and Writing*. (American College Testing, Washington, DC: National Assessment Governing Board, 1993).

The statistical tests determine whether the evidence — based on the data from the groups in the *sample* — is strong enough to conclude that the averages or percentages are really different for those groups in the *population*. If the evidence is strong (i.e., the difference is statistically significant), the report describes the group averages or percentages as being different (e.g., one group performed *higher than* or *lower than* another group) — regardless of whether the sample averages or sample percentages appear to be about the same or not. If the evidence is not sufficiently strong (i.e., the difference is not statistically significant), the averages or percentages are described as being *not significantly different* — again, regardless of whether the sample averages or sample percentages appear to be about the same or widely discrepant. The reader is cautioned to rely on the results of the statistical tests — rather than on the apparent magnitude of the difference between sample averages or percentages — to determine whether those sample differences are likely to represent actual differences between the groups in the population. The statistical tests are discussed in greater detail in Appendix A.

In addition, some of the percentages reported in the text of the report are given quantitative descriptions (e.g., relatively few, about half, almost all, etc.). The descriptive phrases used and the rules used to select them are also described in Appendix A.

How Is This Report Organized?

The *NAEP 1996 Mathematics State Report for New Jersey* is a computer-generated report that describes the mathematics performance of fourth-grade students in New Jersey, the Northeast region, and the nation. A separate report describes additional fourth- and eighth-grade mathematics assessment results for the nation and the states, as well as the national results for grade 12.[11] The *State Report* consists of five sections:

- An **Introduction** provides background information about what was assessed, who was sampled, and how the results are reported.
- **Part One** shows the distribution of mathematics scale score results for fourth-grade students in New Jersey, the Northeast region, and the nation.
- **Part Two** presents mathematics achievement level results for fourth-grade students in New Jersey, the Northeast region, and the nation.

[11] Reese, C.M., K.E. Miller, J. Mazzeo, and J.A. Dossey. *NAEP 1996 Mathematics Report Card*. (Washington, DC: National Center for Education Statistics, 1997).

- **Part Three** relates fourth-grade public school students' mathematics scale scores to contextual information about school characteristics, instruction, and home support for mathematics in New Jersey, the Northeast region, and the nation.

- Several **Appendices** are presented to support the results discussed in the report:

Appendix A	Reporting NAEP 1996 Mathematics Results for New Jersey
Appendix B	NAEP 1996 Mathematics Assessment
Appendix C	Technical Appendix
Appendix D	Setting the Achievement Levels
Appendix E	Teacher Preparation

PART ONE

The Mathematics Scale Score Results for Fourth-Grade Students

The following chapters describe the average mathematics scale scores of fourth-grade students in New Jersey. As described in the Introduction, the NAEP mathematics scale is a composite of the five content strands that comprise the assessment and ranges from 0 to 500.

This part of the report contains two chapters. Chapter 1 compares the overall mathematics performance of public school students in New Jersey to the nation. (Results for the Northeast region are also presented.) Chapter 2 summarizes mathematics performance for subgroups of public school students defined by gender, race/ethnicity, parental education, location of the school, participation in Title I programs and services, and eligibility for the free/reduced-price lunch component of the National School Lunch Program.

CHAPTER 1

Students' Mathematics Scale Score Results

The delivery of education to the millions of school-age students in our country is primarily a function of the states. Therefore, monitoring the performance of students in subjects such as mathematics is a key concern of those policy makers directing educational reform at the state level. Monitoring the mathematics performance of students is also a concern at the national level.

The need to assess the current level of mathematical ability as a mechanism for informing education reform efforts is highlighted by the current National Assessment of Educational Progress (NAEP) in mathematics (as well as the two previous NAEP assessments in mathematics in 1990 and 1992) and the Third International Mathematics and Science Study (TIMSS) conducted in 1994 and 1995 with support from the U.S. Department of Education.[12]

The mathematics community has taken a lead in communicating the importance of mathematics in today's society. With the release of the National Council of Teachers of Mathematics (NCTM) *Curriculum and Evaluation Standards for School Mathematics* in 1989, mathematics educators have accepted the challenges set forth by the national and state policy makers.[13] Based on drafts of the NCTM *Standards*, NAEP developed the 1990 and 1992 mathematics assessments.[14] The framework and specifications for the NAEP 1996 mathematics assessment was refined to better reflect the NCTM *Standards*.[15] Results from the 1996 assessment can be compared to those from the 1990 and 1992 assessments, regardless of the refinement of the framework.

---

[12] The Third International Mathematics and Science Study was conducted in 1994 in the southern hemisphere and in 1995 in the northern hemisphere.

[13] National Council of Teachers of Mathematics. *Curriculum and Evaluation Standards for School Mathematics*. (Reston, VA: NCTM, 1989).

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

[15] National Assessment Governing Board. *Mathematics Framework for the 1996 National Assessment of Educational Progress*. (Washington, DC: National Assessment Governing Board, 1994).

The NAEP 1996 state mathematics assessment at grades 4 and 8 continues the state-level NAEP component begun in 1990 with the NAEP Trial State Assessment (TSA) in mathematics at grade 8, which was followed by the 1992 TSA in mathematics at grades 4 and 8.[16] The current assessment is also the largest with 48 participating jurisdictions.[17] The following results from the NAEP 1996 state mathematics assessment represent a current picture of the mathematics performance of fourth-grade students in New Jersey and the nation.

Table 1.1 shows the distribution of mathematics scale scores for fourth-grade students attending public schools in New Jersey, the Northeast region, and the nation. Results are presented for the 1992 and 1996 assessments and comparisons between the two years are indicated. The 1992 and 1996 mathematics assessments examined the performance of two independent samples of fourth-grade students (i.e., cross-sectional performance). NAEP does not measure the growth of a group of students (i.e., longitudinal performance), but rather collects data for the same grade levels at different points in time.

[16] Based on positive evaluations of the 1990, 1992, and 1994 TSAs, the "Trial" designation has been removed from the 1996 state-level NAEP assessment.

[17] *Jurisdictions* refers to states, territories, the District of Columbia, and Department of Defense Education Activities schools.

1996, Public School Students, Grade 4

The average mathematics scale score in New Jersey was 227. This average was higher than that for the nation (222).[18]

1992 vs. 1996, Public School Students, Grade 4

From 1992 to 1996, the average scale score of students in New Jersey did not change significantly while that of students across the nation increased.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 1.1 — GRADE 4

Distribution of Mathematics Scale Scores for Public School Students

	Average Scale Score	10th Percentile	25th Percentile	50th Percentile	75th Percentile	90th Percentile
1992						
New Jersey	227 (1.5)	187 (3.1)	207 (1.7)	229 (1.5)	249 (1.0)	265 (1.9)
Northeast	224 (2.1)	180 (3.9)	202 (3.1)	226 (3.2)	247 (3.4)	264 (4.3)
Nation	219 (0.8)	176 (1.1)	197 (0.9)	220 (1.0)	241 (1.1)	259 (0.9)
1996						
New Jersey	227 (1.5)	186 (2.8)	207 (1.3)	229 (1.7)	249 (1.6)	266 (1.4)
Northeast	226 (2.6)	181 (4.4)	206 (5.0)	229 (2.4)	250 (1.6)	266 (2.8)
Nation	222 (1.0)>	180 (1.5)	201 (1.3)	224 (1.3)	244 (1.0)	261 (1.0)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

[18] Differences reported as significant are statistically different at the 95 percent confidence level. This means that with 95 percent confidence there is a real difference in the average mathematics scale score between the two populations of interest.

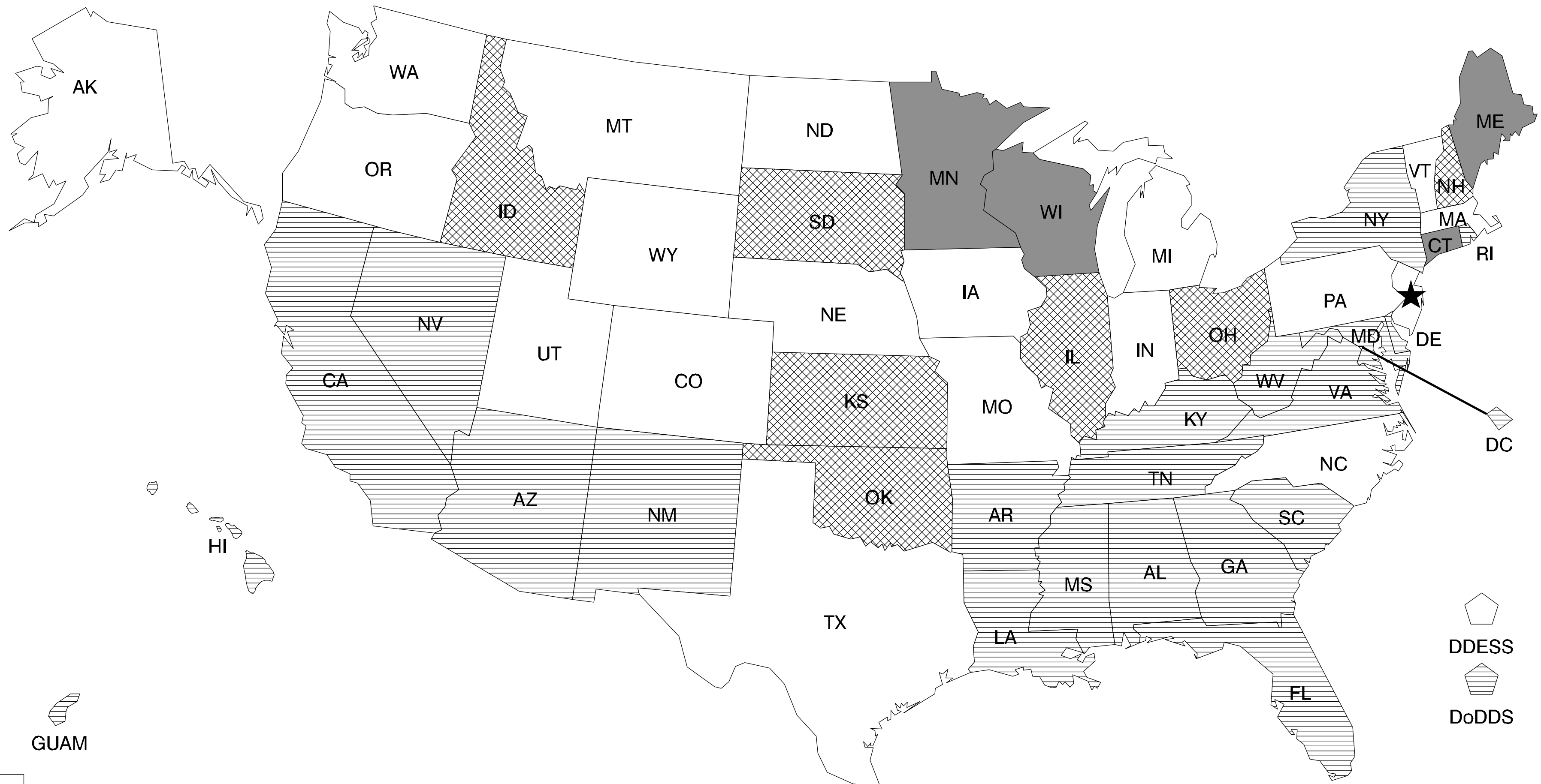
Comparisons Between New Jersey and Other Participating Jurisdictions

The map on the following page provides a method for making appropriate comparisons of the average mathematics scale scores for public school students in New Jersey with those of other jurisdictions participating in the NAEP 1996 mathematics assessment. The different shadings of the states on the map show whether or not the average scale scores of public school students in the other jurisdictions were statistically different from that of public school students in New Jersey ("Target State"). States with horizontal lines have a significantly lower average mathematics scale score than New Jersey and states in gray have a significantly higher average scale score. The unshaded states have average scale scores that did not differ significantly from the average for New Jersey. Several states, those with large crosshatching, did not meet minimum participation rate guidelines established by NCES for the NAEP assessments. A description of the statistical procedures used to produce the data represented in these maps is contained in Appendix A.

The NAEP 1996 State Assessment

Comparisons of Overall Mathematics Scale Scores at Grade 4

New Jersey Public School Students



★ Target state

State has statistically significantly higher average scale score than target state.

State shows no statistically significant difference in average scale score from target state.

State has statistically significantly lower average scale score than target state.

State did not meet minimum participation rate guidelines.

State did not participate.

THE NATION'S REPORT CARD naep 1996 State Assessment

Performance on the NAEP Mathematics Content Strands

The framework and specifications that guided the development of the NAEP mathematics assessments are anchored in broad strands of mathematical content similar to the content standards in the NCTM *Standards*. These content strands are

- Number Sense, Properties, and Operations
- Measurement
- Geometry and Spatial Sense
- Data Analysis, Statistics, and Probability
- Algebra and Functions[19]

Table 1.2 shows the distribution of content strand scale scores for New Jersey, the Northeast region, and the nation. Appendix B describes the five content strands, and Appendix C contains a more extensive discussion of the scaling procedures used to develop the five content-strand scales as well as the composite NAEP mathematics scale.

1996, Public School Students, Grade 4

Students in New Jersey performed higher than students nationwide in number sense, properties, and operations. The performance of students in New Jersey did not differ significantly from that of students nationwide in measurement; geometry and spatial sense; data analysis, statistics, and probability; and algebra and functions.

1992 vs. 1996, Public School Students, Grade 4

The performance of students in New Jersey did not change significantly between 1992 and 1996 in any of the five content strands.

[19] At the fourth-grade level, the Algebra and Functions strand was treated in informal and exploratory ways, often through the study of patterns.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 1.2 — GRADE 4

Distribution of Mathematics Scale Scores for Public School Students by Content Area

	Average Scale Score	10th Percentile	25th Percentile	50th Percentile	75th Percentile	90th Percentile
Number Sense, Properties, and Operations						
1992						
New Jersey	226 (1.6)	183 (2.8)	205 (2.2)	228 (1.9)	249 (1.9)	266 (2.0)
Northeast	221 (2.1)	174 (3.2)	198 (2.5)	223 (3.5)	247 (2.9)	265 (2.6)
Nation	216 (0.9)	170 (1.2)	193 (1.2)	217 (1.1)	240 (0.9)	259 (1.1)
1996						
New Jersey	226 (1.6)	181 (3.1)	204 (2.1)	228 (2.4)	250 (1.3)	268 (1.5)
Northeast	224 (3.2)	175 (5.2)	201 (4.6)	227 (2.6)	250 (1.9)	268 (2.7)
Nation	219 (1.1)	174 (1.1)	197 (0.8)	221 (1.1)	243 (1.3)	262 (1.6)
Measurement						
1992						
New Jersey	231 (1.8)	186 (3.9)	209 (2.7)	233 (1.9)	255 (1.4)	273 (2.2)
Northeast	228 (2.2)	183 (2.7)	205 (1.9)	230 (4.1)	252 (3.1)	270 (4.3)
Nation	223 (0.9)	178 (1.2)	201 (1.3)	225 (0.9)	247 (1.8)	265 (1.0)
1996						
New Jersey	229 (1.9)	183 (3.3)	207 (3.0)	231 (2.1)	253 (1.9)	272 (3.4)
Northeast	227 (2.4)	178 (7.0)	206 (3.8)	230 (2.5)	251 (2.3)	269 (3.5)
Nation	224 (1.2)	178 (2.3)	201 (1.8)	226 (1.5)	248 (1.3)	266 (1.0)
Geometry and Spatial Sense						
1992						
New Jersey	227 (1.3)	189 (2.3)	208 (1.5)	228 (1.6)	246 (1.4)	262 (1.5)
Northeast	225 (2.2)	185 (2.6)	203 (2.9)	225 (3.5)	248 (2.9)	264 (1.2)
Nation	221 (0.7)	181 (1.6)	201 (0.8)	222 (1.1)	243 (1.2)	260 (1.1)
1996						
New Jersey	225 (1.6)	188 (2.7)	206 (2.3)	226 (1.4)	245 (1.3)	261 (2.3)
Northeast	227 (1.8)	186 (3.6)	208 (2.9)	229 (3.1)	248 (0.8)	265 (2.1)
Nation	224 (0.9)	184 (1.1)	205 (0.9)	226 (1.4)	245 (0.8)	261 (1.2)
Data Analysis, Statistics, and Probability						
1992						
New Jersey	226 (1.5)	185 (4.0)	206 (2.1)	228 (1.5)	248 (1.8)	264 (1.9)
Northeast	224 (2.2)	178 (4.1)	202 (2.4)	226 (4.1)	249 (3.7)	265 (3.6)
Nation	219 (1.0)	175 (1.7)	198 (1.2)	221 (1.4)	243 (1.6)	260 (1.6)
1996						
New Jersey	229 (1.8)	190 (2.1)	209 (1.4)	231 (2.0)	249 (1.4)	265 (2.0)
Northeast	228 (3.2)	178 (8.4)	209 (6.1)	232 (2.0)	252 (2.2)	268 (2.0)
Nation	223 (1.4)	180 (1.8)	202 (1.6)	225 (1.4)	246 (1.5)	263 (1.1)
Algebra and Functions						
1992						
New Jersey	226 (1.9)	180 (3.6)	203 (1.8)	227 (1.8)	250 (1.9)	268 (2.2)
Northeast	224 (2.2)	179 (2.4)	202 (2.7)	226 (3.2)	248 (4.0)	266 (2.8)
Nation	218 (0.9)	173 (1.3)	195 (0.9)	219 (1.4)	241 (1.5)	259 (1.8)
1996						
New Jersey	229 (1.5)	187 (1.8)	209 (2.6)	231 (1.9)	251 (1.5)	267 (2.2)
Northeast	231 (2.6)	186 (5.4)	211 (4.6)	234 (2.7)	253 (2.0)	269 (2.1)
Nation	226 (1.2)>	185 (1.8)>	206 (1.3)>	227 (1.7)>	247 (1.5)	264 (1.5)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

CHAPTER 2

Mathematics Scale Score Results by Subpopulations

The previous chapter provided a global view of the mathematics performance of fourth-grade students in New Jersey and the nation. It is also important to examine the average performance of subgroups since past NAEP assessments in mathematics, as well as in other academic subjects, have shown substantial differences among groups defined by gender, racial/ethnic background, parental education, and other demographic characteristics.[20] A key contribution of NAEP to the ongoing conversations on education reform is its ability to monitor the performance of subgroups of students in academic achievement.

The NAEP 1996 state assessment in mathematics provides performance information for subgroups of fourth graders in New Jersey, the Northeast region, and the nation. In addition to the more typical demographic subgroups defined by gender, race/ethnicity, parental education, and location of the school, the 1996 assessment also collected information on student participation in Title I programs and services and eligibility for the free/reduced-price lunch component of the National School Lunch Program.

[20] Campbell, J.R., P.L. Donahue, C.M. Reese, and G.W. Phillips. *NAEP 1994 Reading Report Card for the Nation and the States*. (Washington, DC: National Center for Education Statistics, 1996); Beatty, A.S., C.M. Reese, H.R. Persky, and P. Carr. *NAEP 1994 U.S. History Report Card*. (Washington, DC: National Center for Education Statistics, 1996); Persky, H.R., C.M. Reese, C.Y. O'Sullivan, S. Lazer, J. Moore, and S. Shakrani. *NAEP 1994 Geography Report Card*. (Washington, DC: National Center for Education Statistics, 1996); Campbell, J.R., C.M. Reese, C. O'Sullivan, and J.A. Dossey. *NAEP 1994 Trends in Academic Progress*. (Washington, DC: National Center for Education Statistics, 1996).

A description of the subgroups and how they are defined is presented in Appendix A. The reader is cautioned against making simple or causal inferences related to the performance of various subgroups of students or about the effectiveness of Title I programs. Average performance differences between two groups of students may, in part, be due to socioeconomic or other factors. For example, differences observed among racial/ethnic subgroups are almost certainly associated with a broad range of socioeconomic and educational factors not discussed in this report and possibly not addressed by the NAEP assessment program. Similarly, differences in performance between students participating in Title I programs and those who are not does not account for the initial performance level of the students prior to placement in Title I programs or differences in course content and emphasis between the two groups.

Gender

Consistent with research findings, NAEP mathematics results have shown little difference in the performance of male and female fourth graders.[21] As shown in Table 2.1, the NAEP 1996 state mathematics assessment results for fourth graders in New Jersey are not consistent with those general findings.

1996, Public School Students, Grade 4

The average mathematics scale score of males was higher than that of females in New Jersey; nationwide, however, the performance of males did not differ significantly from that of females.

1992 vs. 1996, Public School Students, Grade 4

From 1992 to 1996 in New Jersey, the average scale score of both males and females did not change significantly.

[21] Frost, L.A., J.S. Hyde, and E. Fennema. "Gender, Mathematics Performance, and Mathematics-Related Attitudes and Affect: A Meta-analytic Synthesis," in *International Journal of Educational Research*, 21, pp. 373-385, 1994; Reese, C.M., K.E. Miller, J. Mazzeo, and J.A. Dossey. *NAEP 1996 Mathematics Report Card*. (Washington, DC: National Center for Education Statistics, 1997).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 2.1 — GRADE 4

Distribution of Mathematics Scale Scores for Public School Students by Gender

		Average Scale Score	10th Percentile	25th Percentile	50th Percentile	75th Percentile	90th Percentile
Male							
1992	New Jersey	228 (1.7)	188 (2.7)	208 (2.3)	230 (1.6)	250 (1.3)	265 (2.6)
	Northeast	226 (2.2)	182 (6.2)	204 (1.4)	229 (2.3)	250 (3.7)	268 (4.8)
	Nation	220 (0.9)	175 (1.3)	198 (0.8)	221 (1.4)	243 (1.4)	261 (1.6)
1996	New Jersey	231 (1.7)	190 (2.8)	210 (1.8)	233 (1.5)	253 (1.5)	270 (2.9)
	Northeast	229 (2.9)	188 (4.3)	209 (4.5)	231 (4.0)	252 (1.9)	270 (1.8)
	Nation	224 (1.2)>	181 (1.8)	202 (2.5)	225 (1.3)	246 (0.8)	264 (1.6)
Female							
1992	New Jersey	226 (1.6)	186 (1.9)	206 (2.1)	228 (2.0)	248 (2.0)	263 (3.7)
	Northeast	221 (2.8)	177 (3.3)	201 (5.5)	223 (4.7)	244 (3.8)	261 (3.2)
	Nation	218 (1.1)	176 (1.2)	197 (2.2)	219 (1.1)	239 (2.2)	257 (1.1)
1996	New Jersey	223 (1.7)	184 (1.6)	204 (1.9)	225 (1.7)	244 (2.3)	260 (2.9)
	Northeast	224 (2.8)	174 (7.0)	203 (9.3)	229 (3.5)	248 (1.9)	263 (2.4)
	Nation	221 (1.1)	180 (1.9)	201 (1.6)	223 (1.3)	242 (0.9)	259 (1.0)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

Race/Ethnicity

As part of the background questions administered with the NAEP 1996 mathematics assessment, students were asked to identify the racial/ethnic subgroup that best describes them. The five mutually exclusive categories were White, Black, Hispanic, Asian or Pacific Islander, and American Indian or Alaskan Native.

Research over past decades has shown that racial/ethnic differences exist in mathematics performance, and findings from previous NAEP assessments are consistent with this body of research.[22] However, when interpreting differences in subgroup performance, confounding factors related to socioeconomic status, home environment, and educational opportunities available to students need to be considered.[23] The distribution of fourth-grade mathematics scale scores for New Jersey, the Northeast region, and the nation are shown in Table 2.2 for White, Black, Hispanic, and Asian/Pacific Islander students.[24]

1996, Public School Students, Grade 4

White students in New Jersey demonstrated an average mathematics scale score that was higher than that of Black and Hispanic students but was lower than that of Asian/Pacific Islander students.

1992 vs. 1996, Public School Students, Grade 4

From 1992 to 1996, there was no significant change in the average scale score of White, Black, Hispanic, or Asian/Pacific Islander students in New Jersey.

---

[22] Campbell, J.R., C.M. Reese, C. O'Sullivan, and J.A. Dossey. *NAEP 1994 Trends in Academic Progress*. (Washington, DC: National Center for Education Statistics, 1996); Reese, C.M., K.E. Miller, J. Mazzeo, and J.A. Dossey. *NAEP 1996 Mathematics Report Card*. (Washington, DC: National Center for Education Statistics, 1997).

[23] McKenzie, F.D. "Educational Strategies for the 1990s," in *The State of Black America 1991*. (New York, NY: National Urban League, Inc., 1991).

[24] Results are reported for racial/ethnic subgroups meeting established sample size requirements (see Appendix A).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 2.2 — GRADE 4

Distribution of Mathematics Scale Scores for Public School Students by Race/Ethnicity

		Average Scale Score	10th Percentile	25th Percentile	50th Percentile	75th Percentile	90th Percentile
White							
1992	New Jersey	237 (1.2)	203 (2.0)	220 (1.6)	238 (1.0)	254 (1.4)	268 (2.2)
	Northeast	233 (2.3)	195 (2.7)	213 (4.1)	235 (3.0)	253 (2.3)	269 (3.5)
	Nation	227 (1.0)	189 (1.6)	208 (1.2)	229 (0.9)	247 (1.2)	263 (1.2)
1996	New Jersey	239 (1.2)	208 (1.7)	223 (1.4)	241 (1.5)	256 (1.1)	271 (2.3)
	Northeast	237 (1.9)	201 (5.2)	220 (1.8)	237 (1.5)	256 (1.7)	270 (3.6)
	Nation	231 (1.1)>	195 (1.2)	213 (2.0)	232 (1.0)	250 (0.9)	266 (1.0)
Black							
1992	New Jersey	199 (2.5)	164 (4.9)	180 (3.4)	200 (2.9)	217 (3.6)	231 (3.6)
	Northeast	195 (3.1)	156 (13.6)	174 (3.5)	196 (6.6)	215 (2.5)	234 (5.5)
	Nation	192 (1.4)	155 (2.6)	173 (2.5)	193 (1.7)	211 (1.7)	228 (2.7)
1996	New Jersey	204 (2.4)	170 (5.2)	186 (3.9)	205 (2.2)	222 (2.5)	236 (3.9)
	Northeast	193 (9.1)!	147 (10.5)!	169 (10.6)!	191 (8.8)!	219 (5.9)!	241 (11.8)!
	Nation	200 (2.4)	163 (2.3)	181 (1.9)	200 (2.8)	220 (3.2)	238 (3.8)
Hispanic							
1992	New Jersey	206 (2.6)	167 (4.2)	186 (3.2)	208 (2.8)	227 (2.5)	243 (4.2)
	Northeast	201 (3.1)	166 (8.4)	185 (3.4)	201 (3.3)	219 (7.0)	236 (7.3)
	Nation	201 (1.5)	163 (3.1)	181 (1.8)	201 (1.4)	221 (1.7)	239 (2.6)
1996	New Jersey	206 (2.9)	169 (3.7)	187 (4.6)	206 (3.6)	226 (4.5)	242 (1.5)
	Northeast	212 (4.8)	171 (11.4)	193 (9.8)	213 (4.2)	233 (7.2)	251 (6.9)
	Nation	205 (2.2)	163 (3.7)	184 (2.9)	206 (2.1)	227 (2.1)	244 (0.8)
Asian/Pacific Islander							
1992	New Jersey	241 (2.9)	203 (6.0)	223 (4.9)	241 (4.1)	261 (8.1)	275 (2.5)
	Northeast	*** (**.*)	*** (**.*)	*** (**.*)	*** (**.*)	*** (**.*)	*** (**.*)
	Nation	233 (2.5)	193 (5.5)	215 (6.1)	234 (3.0)	254 (5.2)	270 (5.1)
1996	New Jersey	248 (2.7)	218 (23.9)	230 (6.5)	248 (1.6)	266 (3.8)	280 (2.8)
	Northeast	*** (**.*)	*** (**.*)	*** (**.*)	*** (**.*)	*** (**.*)	*** (**.*)
	Nation	231 (4.6)	192 (9.3)	212 (6.0)	232 (7.3)	248 (6.0)	273 (8.9)

The NAEP mathematics scale ranges from 0 to 500. Results are reported for racial/ethnic subgroups meeting established sample size requirements (see Appendix A). The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic. *** Sample size is insufficient to permit a reliable estimate.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

Students' Reports of Parents' Highest Education Level

Students were asked to indicate the level of education completed by each parent. Four levels of education were identified: did not finish high school, graduated high school, some education after high school, and graduated college. A choice of "I don't know" was also available. For this analysis the highest education level reported for either parent was used.

In general, results show that with each increment in reported parental education, student performance increases significantly. In reviewing these results, it is important to note that nationally, approximately one third of fourth graders did not know the level of education that either of their parents had completed. For public school students in New Jersey, this percentage was 32 percent. Despite the fact that some research has questioned the accuracy of student-reported data from similar groups of students,[25] past NAEP assessments in mathematics, as well as other subject areas, have found that student-reported level of parental education exhibits a consistent positive relationship with student performance on the assessments.[26] Other research has also replicated NAEP findings.[27]

Table 2.3 shows the results for fourth-grade public school students who reported that neither parent graduated from high school, at least one parent graduated from high school, at least one parent had some education after high school, at least one parent graduated from college, or that they did not know their parents' highest education level. The following discussion pertains to those students who reported knowing the educational level of one or both parents.

1996, Public School Students, Grade 4

The average mathematics scale score of students in New Jersey who reported that neither parent graduated from high school was lower than that of students who reported that at least one parent graduated from high school, at least one parent had some education after high school, or at least one parent graduated from college.

1992 vs. 1996, Public School Students, Grade 4

The average scale score of students in New Jersey who reported that neither parent graduated from high school, at least one parent graduated from high school, at least one parent had some education after high school, or at least one parent graduated from college did not change significantly between 1992 and 1996.

---

[25] Looker, E.D. "Accuracy of Proxy Reports of Parental Status Characteristics," in *Sociology of Education*, 62(4), pp. 257-276, 1989.

[26] Reese, C.M., K.E. Miller, J. Mazzeo, and J.A. Dossey. *NAEP 1996 Mathematics Report Card*. (Washington, DC: National Center for Education Statistics, 1997); Campbell, J.R., P.L. Donahue, C.M. Reese, and G.W. Phillips. *NAEP 1994 Reading Report Card for the Nation and the States*. (Washington, DC: National Center for Education Statistics, 1996); Beatty, A.S., C.M. Reese, H.R. Persky, and P. Carr. *NAEP 1994 U.S. History Report Card*. (Washington, DC: National Center for Education Statistics, 1996); Persky, H.R., C.M. Reese, C.Y. O'Sullivan, S. Lazer, J. Moore, and S. Shakrani. *NAEP 1994 Geography Report Card*. (Washington, DC: National Center for Education Statistics, 1996).

[27] National Education Longitudinal Study. *National Education Longitudinal Study of 1988: Base Year Student Survey*. (Washington, DC: National Center for Education Statistics, 1995).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 2.3 — GRADE 4

Distribution of Mathematics Scale Scores by Public School Students' Reports of Parents' Highest Education Level

		Average Scale Score	10th Percentile	25th Percentile	50th Percentile	75th Percentile	90th Percentile
Did not finish high school							
1992	New Jersey	211 (3.7)	172 (15.7)	193 (8.1)	213 (6.4)	230 (15.1)	246 (9.2)
	Northeast	*** (**.*)	*** (**.*)	*** (**.*)	*** (**.*)	*** (**.*)	*** (**.*)
	Nation	204 (2.6)	166 (5.1)	186 (3.0)	205 (3.9)	224 (4.8)	241 (5.6)
1996	New Jersey	207 (4.6)	168 (10.3)	185 (7.4)	204 (9.1)	230 (3.8)	245 (4.6)
	Northeast	*** (**.*)	*** (**.*)	*** (**.*)	*** (**.*)	*** (**.*)	*** (**.*)
	Nation	205 (2.5)	170 (8.6)	185 (5.6)	205 (1.8)	224 (4.8)	240 (4.6)
Graduated from high school							
1992	New Jersey	220 (2.0)	184 (5.0)	201 (3.8)	221 (1.8)	240 (2.9)	254 (6.6)
	Northeast	216 (5.1)	177 (5.7)	197 (4.1)	217 (5.9)	234 (5.6)	254 (6.6)
	Nation	214 (1.6)	174 (3.0)	193 (2.8)	216 (1.5)	234 (1.5)	252 (2.7)
1996	New Jersey	222 (3.1)	184 (5.5)	206 (4.3)	224 (3.4)	243 (4.9)	257 (1.9)
	Northeast	224 (4.3)	177 (4.4)	201 (9.2)	228 (3.4)	250 (5.0)	263 (3.1)
	Nation	218 (1.7)	178 (3.7)	198 (1.6)	220 (2.1)	240 (1.6)	255 (2.0)
Some education after HS							
1992	New Jersey	231 (2.7)	196 (3.4)	214 (2.5)	233 (3.8)	250 (5.9)	263 (6.0)
	Northeast	230 (4.6)	185 (14.9)	210 (12.6)	235 (4.0)	251 (4.7)	268 (5.2)
	Nation	224 (1.7)	181 (1.5)	204 (3.8)	228 (2.0)	246 (2.3)	260 (2.6)
1996	New Jersey	233 (2.2)	198 (3.0)	219 (2.8)	234 (6.5)	249 (3.0)	260 (5.5)
	Northeast	237 (3.3)	196 (2.2)	220 (5.4)	237 (4.6)	257 (5.2)	274 (3.8)
	Nation	232 (1.7)>	197 (2.8)>	216 (6.1)	233 (1.1)	250 (2.7)	265 (3.6)
Graduated from college							
1992	New Jersey	235 (1.7)	196 (2.8)	216 (1.9)	237 (1.5)	256 (2.2)	270 (1.7)
	Northeast	232 (3.0)	185 (3.6)	210 (3.4)	235 (5.2)	256 (4.1)	272 (3.4)
	Nation	226 (1.1)	181 (3.2)	205 (1.5)	228 (2.1)	248 (1.5)	266 (2.0)
1996	New Jersey	234 (2.1)	191 (2.2)	213 (2.4)	236 (2.6)	256 (1.8)	273 (4.4)
	Northeast	232 (4.4)	185 (10.7)	213 (4.8)	237 (2.4)	257 (2.6)	271 (3.0)
	Nation	230 (1.6)	188 (2.9)	210 (1.2)	232 (2.1)	251 (1.4)	268 (1.6)
I don't know							
1992	New Jersey	219 (2.0)	177 (3.0)	198 (3.6)	221 (3.3)	241 (1.2)	256 (1.7)
	Northeast	217 (2.4)	177 (2.3)	197 (2.9)	217 (3.6)	239 (3.0)	256 (2.9)
	Nation	213 (0.9)	173 (1.4)	193 (1.0)	214 (1.3)	235 (1.5)	252 (1.6)
1996	New Jersey	221 (1.7)	182 (3.4)	201 (4.1)	222 (1.4)	242 (1.6)	257 (1.4)
	Northeast	221 (3.5)	178 (11.3)	203 (6.6)	223 (5.0)	241 (2.5)	258 (8.1)
	Nation	216 (1.5)	175 (4.6)	196 (2.3)	218 (1.3)	238 (1.2)	255 (2.4)

The NAEP mathematics scale ranges from 0 to 500. Results are reported for parental education subgroups meeting established sample size requirements (see Appendix A). The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. *** Sample size is insufficient to permit a reliable estimate.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

Title I Participation

The Improving America's Schools Act of 1994 (P.L. 103-382) reauthorized the Elementary and Secondary Education Act of 1965 (ESEA). Title I Part A of the ESEA provides local education agencies with financial assistance to meet the educational needs of children who are failing or most at risk of failing.[28] Title I programs are designed to help disadvantaged students meet challenging academic performance standards. Through Title I, schools are assisted in improving teaching and learning and in providing students with opportunities to acquire the knowledge and skills outlined in their state's content and performance standards. For high poverty Title I schools, all children in the school may benefit through participation in schoolwide programs. Title I funding supports state and local education reform efforts and promotes coordinating of resources to improve education for all students.

NAEP first collected student-level information on participation in Title I programs in 1994. Therefore, results comparing the performance of participating and nonparticipating students are not available for previous NAEP mathematics assessments. The NAEP program will continue to monitor the performance of Title I program participants in future assessments. The Title I information collected by NAEP refers to current participation in Title I services. Students who participated in such services in the past but do not currently receive services are not identified as Title I participants. Differences between students who receive Title I services and those who do not should not be viewed as an evaluation of Title I programs. Typically, Title I services are intended for students who score poorly on assessments. To properly evaluate Title I programs, the performance of students participating in such programs must be monitored over time and their progress must be assessed.[29]

Table 2.4 presents results for fourth-grade students by Title I participation.

1996, Public School Students, Grade 4

The average mathematics scale score of students in New Jersey who received Title I services (195) was not significantly different from that of students nationwide (200). The average scale score of New Jersey students who did not receive Title I services (233) was not significantly different from the national average (229). The average scale score of New Jersey students who received Title I services was lower than that of students who did not.

[28] U.S. Department of Education, Office of Elementary and Secondary Compensatory Education Programs. *Improving Basic Programs Operated by Local Education Agencies*. (Washington, DC: U.S. Department of Education, 1996).

[29] For a study of mathematics performance of Title I students in 1991-1992, see U.S. Department of Education, *PROSPECTS: The Congressionally Mandated Study of Educational Growth and Opportunity, Interim Report: Language, Minority and Limited English Proficient Students*. (Washington, DC: U.S. Department of Education, 1995).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 2.4 — GRADE 4

Distribution of Mathematics Scale Scores for Public School Students by Title I Participation

		Average Scale Score	10th Percentile	25th Percentile	50th Percentile	75th Percentile	90th Percentile
GRADE 4							
Participated							
1996	New Jersey	195 (3.2)	161 (6.3)	178 (3.2)	195 (4.0)	213 (1.9)	227 (4.2)
	Northeast	197 (3.4)	157 (11.3)	179 (3.6)	198 (3.6)	218 (3.0)	231 (6.1)
	Nation	200 (1.9)	164 (2.1)	181 (1.9)	200 (2.5)	219 (1.6)	235 (2.6)
Did not participate							
1996	New Jersey	233 (1.4)	195 (1.3)	214 (2.1)	234 (2.0)	252 (0.9)	268 (1.8)
	Northeast	233 (2.4)	195 (3.7)	216 (3.4)	236 (1.8)	254 (2.4)	269 (2.2)
	Nation	229 (1.1)	191 (1.2)	211 (1.3)	231 (1.1)	249 (1.6)	265 (1.2)

The NAEP mathematics scale ranges from 0 to 500. Results are reported for students participating in Title I programs only if established sample size requirements are met (see Appendix A). The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

Free/Reduced-Price Lunch Program Eligibility

The free/reduced-price lunch component of the National School Lunch Program (NSLP), offered through the U.S. Department of Agriculture (USDA), is designed to ensure that children near or below the poverty line receive nourishing meals.[30] Eligibility for the free/reduced-price lunch program is included as an indicator of poverty. The program is available to public schools, nonprofit private schools, and residential child care institutions. Eligibility for free or reduced-price meals is determined through the USDA's Income Eligibility Guidelines.

NAEP first collected information on student-level eligibility for the federally funded NSLP in 1996. Although results cannot be presented for previous NAEP mathematics assessments, the NAEP program will continue to monitor the performance of these students in future assessments. Table 2.5 shows the results for fourth graders based on their eligibility for this program.

1996, Public School Students, Grade 4

The average mathematics scale score of students in New Jersey who were eligible for free or reduced-price lunch (206) was not significantly different from that of students nationwide (207). The average scale score of New Jersey students who were not eligible for free or reduced-price lunch (238) was higher than the national average (231). The average scale score of New Jersey students who were eligible for free or reduced-price lunch was lower than that of students who were not.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 2.5 — GRADE 4

Distribution of Mathematics Scale Scores for Public School Students by Free/Reduced-Price Lunch Eligibility

		Average Scale Score	10th Percentile	25th Percentile	50th Percentile	75th Percentile	90th Percentile
GRADE 4							
Eligible							
1996	New Jersey	206 (2.2)	171 (5.6)	188 (3.6)	207 (2.0)	225 (2.3)	241 (1.4)
	Northeast	205 (5.9)	157 (8.6)	180 (6.6)	206 (10.6)	233 (4.2)	250 (3.1)
	Nation	207 (2.0)	167 (3.6)	185 (2.4)	207 (2.3)	230 (2.0)	246 (1.2)
Not eligible							
1996	New Jersey	238 (1.4)	203 (2.4)	221 (1.3)	239 (1.3)	256 (1.2)	271 (4.0)
	Northeast	237 (1.8)	204 (4.1)	220 (2.1)	238 (1.7)	256 (2.8)	271 (3.2)
	Nation	231 (1.1)	195 (1.6)	213 (1.6)	231 (1.3)	249 (1.5)	266 (1.6)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

[30] U.S. General Services Administration. *Catalog of Federal Domestic Assistance*. (Washington, DC: Executive Office of the President, Office of Management and Budget, 1995).

Type of Location

For the purpose of reporting, schools that participated in the assessment were classified into three mutually exclusive types of location — Central City, Urban Fringe/Large Town, and Rural/Small Town. These classifications are based on geographic characteristics of the schools' locations and are determined by Census Bureau definitions of metropolitan statistical areas (MSAs), population size, and density. These categories indicate the geographic locations of schools and are not intended to indicate or imply social or economic meanings for location types.[31] (The type of location classification is described in Appendix A.)

Table 2.6 presents fourth-grade results according to the location type of the schools that students attended for New Jersey and the nation.[32]

1996, Public School Students, Grade 4

The average mathematics scale score of students attending schools in central cities in New Jersey was lower than that of students in urban fringes/large towns.

1992 vs. 1996, Public School Students, Grade 4

From 1992 to 1996, there was no significant change in the average scale score of students attending schools in central cities or urban fringes/large towns in New Jersey.

---

[31] In past NAEP reports, a type of community variable that combined community size with a school-level socioeconomic indicator was reported. Due to the problematic nature of this variable, NAEP currently reports results by Census-based descriptors.

[32] Results are reported for type of location subgroups meeting established sample size requirements (see Appendix A).

TABLE 2.6 — GRADE 4

Distribution of Mathematics Scale Scores for Public School Students by Type of Location

		Average Scale Score	10th Percentile	25th Percentile	50th Percentile	75th Percentile	90th Percentile
Central city							
1992	New Jersey	202 (4.1)	163 (4.6)	179 (6.3)	202 (6.9)	222 (4.8)	243 (6.5)
	Nation	212 (1.6)	168 (2.2)	189 (3.1)	212 (2.1)	235 (2.2)	253 (2.0)
1996	New Jersey	213 (6.1)!	172 (6.7)!	190 (5.7)!	213 (6.1)!	236 (10.7)!	254 (7.2)!
	Nation	214 (2.8)	170 (3.5)	190 (3.7)	215 (3.5)	238 (2.9)	256 (2.5)
Urban fringe/Large town							
1992	New Jersey	231 (1.6)	194 (2.3)	213 (1.6)	233 (1.9)	251 (2.2)	267 (2.2)
	Nation	225 (1.6)	184 (4.0)	205 (1.8)	227 (2.4)	247 (1.7)	264 (1.4)
1996	New Jersey	229 (1.9)	190 (3.7)	210 (3.5)	231 (2.0)	250 (1.2)	266 (1.9)
	Nation	228 (1.7)	188 (3.3)	209 (1.8)	229 (1.2)	248 (2.1)	265 (1.7)

The NAEP mathematics scale ranges from 0 to 500. Results are reported for type of location subgroups meeting established sample size requirements (see Appendix A). The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. Characteristics of the school sample do not permit reliable regional results for type of location. ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

PART TWO

The Mathematics Achievement Level Results for Fourth-Grade Students

While providing information about what students can do in mathematics is essential for understanding the current state of mathematics performance, knowing what students *can do* is made even more relevant by also looking at what students *should be able to do*. For that reason, the National Assessment Governing Board (NAGB) has provided NAEP with achievement levels in mathematics that set standards for performance in mathematics at grades 4, 8, and 12.

This part of the report presents results using the student achievement levels as authorized by the NAEP legislation and adopted by NAGB.[33] The achievement levels are based on collective judgments, gathered from a broadly representative panel of teachers, education specialists, and members of the general public, about what students should know and be able to do relative to a body of content reflected in the NAEP mathematics frameworks. For reporting purposes, the achievement level cutscores are placed on the traditional NAEP scale. For each grade, the results divide the scale into four ranges — *Basic, Proficient, Advanced*, and the region below *Basic*.

Initiated in 1990, the levels have been used to report the national and state results in mathematics in 1990 and 1992, as well as in other subjects such as reading, U.S. history, and geography. The mathematics achievement levels were developed by American College Testing (ACT) under contract with NAGB. While setting student achievement levels on NAEP is relatively new and developing, the achievement levels are consistent with recent education reform efforts. Some state and local jurisdictions are also developing standards and reporting their test results using them.[34]

---

[33] P.L. 103-382. Improving America's Schools Act of 1994.

[34] States such as Kentucky, Maryland, Colorado, Connecticut, and North Carolina all have standard-setting initiatives resulting in student achievement levels.

Despite the commitment to standards-based reporting of NAEP results, the transition is incomplete. There have been some critical reviews and congressionally mandated evaluations that cast doubt on the interpretability of achievement levels and also on the applicability of the underlying technical methodology used to develop them. These studies were conducted by the General Accounting Office (GAO)[35] and the National Academy of Education (NAE).[36] Their findings question, for example, the application of the Angoff method for large-scale assessments like NAEP, given the significant modifications required to accommodate the complexity of the NAEP item structure and the multiple cutpoints. They conclude that discretion should be used in making particular inferences about what students at each level actually know and can do. In addition, there were concerns that the proportion of students at certain levels, but particularly at the *Advanced* levels, may be underestimated.

On the other hand, the Angoff procedure is the most widely documented, researched, and frequently used method in the standard-setting field. Many well-known experts support the use of a modified-Angoff method on NAEP. Several critics of the NAE studies,[37] for example, have reaffirmed the integrity of the process employed by NAGB and have concluded that the weight of the empirical evidence presented does not support the NAE's conclusions about achievement levels or the use of the modified-Angoff process. In addition, the Council of Chief State School Officers' advisory panel of state assessment directors, fully aware of the NAE's conclusions, supported the use of the achievement levels to report NAEP results.[38]

Taken together, the results of the various studies suggest the need for further research and development. A standard-setting conference was held in the fall of 1994, jointly sponsored by NCES and NAGB. The proceedings, which were recently released, show the variety of approaches which can be used to achieve similar goals and the general lack of agreement on which approach may constitute the best way.[39]

---

[35] General Accounting Office. *Educational Achievement Standards: NAGB's Approach Yields Misleading Interpretations*. (Washington, DC: General Accounting Office, 1993).

[36] National Academy of Education. *Setting Performance Standards for Student Achievement*. (Stanford, CA: National Academy of Education, 1993).

[37] American College Testing. *Technical Report on Setting Achievement Levels on the 1992 National Assessment of Educational Progress in Mathematics, Reading, and Writing*. (Washington, DC: National Assessment Governing Board, 1993); G. Cizek. *Reactions to the National Academy of Education Report*. (Washington, DC: National Assessment Governing Board, 1993); M. Kane. *Comments on the NAE Evaluation of the NAGB Achievement Levels*. (Washington, DC: National Assessment Governing Board, 1993).

[38] Education Information Advisory Committee of the Council of Chief State School Officers. *A Resolution of the Education Information Advisory Committee*. (Alexandria, VA, 1994).

[39] National Assessment Governing Board (NAGB) and National Center for Education Statistics (NCES). *Joint Conference on Standard Setting for Large Scale Assessments*. (Washington, DC: U.S. Government Printing Office, 1995).

In summary, the student achievement levels in this report have been developed carefully and responsibly, and have been subject to refinements and revisions in procedures as new technologies have become available. However, standards-based reporting for NAEP data is still in transition. The NAEP legislation states that the student achievement levels shall be ". . . developed through a national consensus approach, . . . used on a developmental basis, . . . and updated as appropriate." It requires that the developmental status of achievement levels be clearly stated in NAEP reports. Upon review of the available information, the Commissioner of Education Statistics has judged that the achievement levels are in a developmental status. However, the Commissioner and NAGB also believe that the achievement levels are useful and valuable in reporting on the educational achievement of American students.

Part Two of this report focuses on results of the NAEP 1996 state assessment program in mathematics in terms of the NAGB achievement levels. Chapter 3 provides an overview of the achievement level descriptors. In addition, the percentages of public school students in New Jersey, the Northeast region, and the nation who performed at or above each of the achievement levels are presented. Chapter 4 expands on these results by presenting achievement level data for subgroups defined by gender, race/ethnicity, parental education, location of the school, participation in Title I services and programs, and eligibility for the free/reduced-price lunch component of the National School Lunch Program.

CHAPTER 3

Students' Mathematics Achievement Level Results

Achievement levels are based on collective judgments, gathered from a broadly representative panel of teachers, education specialists, and members of the general public, about what students should know and be able to do relative to the body of content reflected in the NAEP mathematics framework (see Appendix B for a description of the framework). These judgments translate into specific points on the NAEP scale that identify boundaries between levels of achievement — *Basic, Proficient*, and *Advanced* — for each grade. Performance at the *Basic* level denotes partial mastery of the knowledge and skills that are fundamental for proficient work. Performance at the *Proficient* level, represents solid academic performance. Students reaching this level demonstrate competency over challenging subject matter. Performance at the *Advanced* level signifies superior performance beyond proficient grade-level mastery. In this report, the percentage of students at or above the three achievement levels, as well as the percentage of students below *Basic*, is presented for fourth-grade students in New Jersey, the Northeast region, and the nation.

Description of NAEP Mathematics Achievement Levels

The achievement levels for the NAEP mathematics assessments were first set in 1990 and slightly revised following the 1992 mathematics assessment. Appendix D briefly describes the process of gathering expert judgments about *Basic, Proficient*, and *Advanced* performance — as defined by NAGB policy — on each mathematics question. The appendix also discusses procedures for combining the various judgments on the various questions and mapping them onto the NAEP mathematics scale. The result of the achievement level setting process is a set of scale score cutpoints used to classify students as *Basic, Proficient*, or *Advanced*. (Separate cutpoints are defined for each grade.) The three mathematics achievement levels for grade 4 are elaborated on in Figure 3.1, and examples of questions appropriate at each achievement level are also provided. It should be noted that constructed-response questions in the assessment occur at all levels of mathematics achievement.

THE NATION'S REPORT CARD naep 1996 State Assessment

FIGURE 3.1

Mathematics Achievement Levels

GRADE 4

NAEP mathematics content strands: (1) Number Sense, Properties, and Operations; (2) Measurement; (3) Geometry and Spatial Sense; (4) Data Analysis, Statistics, and Probability; (5) Algebra and Functions. (Note: At the fourth-grade level, algebra and functions are treated in informal and exploratory ways, often through the study of patterns.)

Skills are cumulative across levels — from *Basic* to *Proficient* to *Advanced*.

BASIC LEVEL	**Fourth-grade students performing at the *Basic* level *should show some evidence of understanding the mathematical concepts and procedures in the five NAEP content strands*. In relation to the NAEP mathematics scale, *Basic*-level achievement for fourth grade is defined by scale scores at or above 214.**

Specifically, fourth graders performing at the *Basic* level should be able to estimate and use basic facts to perform simple computations with whole numbers; show some understanding of fractions and decimals; and solve simple real-world problems in all NAEP content strands. Students at this level should be able to use — though not always accurately — four-function calculators, rulers, and geometric shapes. Their written responses are often minimal and presented without supporting information.

PROFICIENT LEVEL	**Fourth-grade students performing at the *Proficient* level *should consistently apply integrated procedural knowledge and conceptual understanding to problem solving in the five NAEP content strands*. In relation to the NAEP mathematics scale, *Proficient*-level achievement for fourth grade is defined by scale scores at or above 249.**

Specifically, fourth graders performing at the *Proficient* level should be able to use whole numbers to estimate, compute, and determine whether results are reasonable. They should have a conceptual understanding of fractions and decimals; be able to solve real-world problems in all NAEP content strands; and use four-function calculators, rulers, and geometric shapes appropriately. Students at the *Proficient* level should employ problem-solving strategies such as identifying and using appropriate information. Their written solutions should be organized and presented both with supporting information and explanations of how they were achieved.

ADVANCED LEVEL	**Fourth-grade students performing at the *Advanced* level *should apply integrated procedural knowledge and conceptual understanding to complex and nonroutine real-world problem solving in the five NAEP content strands*. In relation to the NAEP scale, *Advanced*-level achievement for fourth grade is defined by scale scores at or above 282.**

Specifically, fourth graders performing at the *Advanced* level should be able to solve complex and nonroutine real-world problems in all NAEP content strands. They should display mastery in the use of four-function calculators, rulers, and geometric shapes. These students are expected to draw logical conclusions and justify answers and solution processes by explaining why, as well as how, they were achieved. They should go beyond the obvious in their interpretations and be able to communicate their thoughts clearly and concisely.

THE NATION'S REPORT CARD naep 1996 State Assessment

FIGURE 3.1 (continued)

Mathematics Achievement Levels

Grade 4 Basic-Level Example Item

Refer to the rectangle below. (NOTE: Size reduced from original.)

Use your centimeter ruler to make the following measurement to the nearest centimeter.

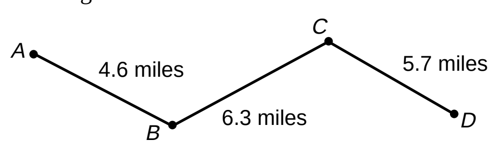
What is the length in centimeters of one of the longer sides of the rectangle?

Answer: (8 centimeters)

1992 Percent Correct	
Nation	50 (1.6)

Grade 4 Proficient-Level Example Item

Carol wanted to estimate the distance from *A* to *D* along the path shown on the map below. She correctly rounded each of the given distances to the nearest mile and then added them. Which of the following sums could be hers?



A. 4 + 6 + 5 = 15
B. 5 + 6 + 5 = 16
*C. 5 + 6 + 6 = 17
D. 5 + 7 + 6 = 18

1992 Percent Correct	
Nation	25 (1.7)

Grade 4 Advanced-Level Example Item

If □ represents the number of newspapers that Lee delivers each day, which of the following represents the total number of newspapers that Lee delivers in 5 days?

A. 5 + □
*B. 5 x □
C. □ ÷ 5
D. (□ + □) x 5

1992 Percent Correct	
Nation	48 (1.4)

Table 3.1 indicates the percentage of fourth-grade public school students at or above each achievement level, as well as the percentage of students below the *Basic* level.

1996, Public School Students, Grade 4

In New Jersey, 25 percent of students performed at or above the *Proficient* level. This percentage was larger than that of students nationwide (20 percent).

1992 vs. 1996, Public School Students, Grade 4

From 1992 to 1996, there was no significant change in the percentage of New Jersey students who attained the *Proficient* level (25 percent in 1992 and 25 percent in 1996).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 3.1 — GRADE 4

Percentage of Public School Students Attaining Mathematics Achievement Levels

	Advanced	At or Above Proficient	At or Above Basic	Below Basic
1992				
New Jersey	2 (0.6)	25 (1.5)	68 (2.1)	32 (2.1)
Northeast	3 (0.7)	23 (2.8)	62 (2.9)	38 (2.9)
Nation	2 (0.3)	17 (1.1)	57 (1.2)	43 (1.2)
1996				
New Jersey	3 (0.7)	25 (1.7)	68 (2.1)	32 (2.1)
Northeast	3 (1.0)	26 (1.6)	68 (3.5)	32 (3.5)
Nation	2 (0.3)	20 (1.0)	62 (1.4)>	38 (1.4)<

The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

CHAPTER 4

Mathematics Achievement Level Results by Subpopulations

As discussed in Chapter 2 of this report, NAEP assessments have repeatedly shown differences in performance for subpopulations of students. This chapter presents achievement level results for New Jersey, the Northeast region, and the nation for subgroups of public school students defined by gender, race/ethnicity, parental education, type of location of the students' schools, participation in Title I programs, and eligibility for the free/reduced-price lunch component of the National School Lunch Program. (A description of the subgroups and how they are defined is presented in Appendix A.)

As stated in Part One, the reader is cautioned against using NAEP results to make simple or causal inferences related to subgroup membership or the effectiveness of Title I programs.

Gender

Table 4.1 provides the achievement level results by gender for fourth-grade public school students.

1996, Public School Students, Grade 4

The percentage of males in New Jersey who performed at or above the *Proficient* level (30 percent) was larger than that of females (20 percent).

1992 vs. 1996, Public School Students, Grade 4

In New Jersey from 1992 to 1996, the percentage of both males and females who attained the *Proficient* level did not change significantly.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 4.1 — GRADE 4

Percentage of Public School Students Attaining Mathematics Achievement Levels by Gender

		Advanced	At or Above Proficient	At or Above Basic	Below Basic
Male					
1992	New Jersey	3 (0.9)	26 (1.9)	69 (2.5)	31 (2.5)
	Northeast	3 (1.0)	26 (3.2)	65 (2.7)	35 (2.7)
	Nation	2 (0.4)	19 (1.2)	59 (1.3)	41 (1.3)
1996	New Jersey	3 (1.0)	30 (2.6)	72 (2.7)	28 (2.7)
	Northeast	4 (1.5)	28 (2.1)	69 (4.0)	31 (4.0)
	Nation	3 (0.5)	22 (1.2)	63 (1.8)	37 (1.8)
Female					
1992	New Jersey	2 (0.6)	23 (2.0)	67 (2.4)	33 (2.4)
	Northeast	2 (0.8)	19 (3.4)	60 (4.5)	40 (4.5)
	Nation	1 (0.3)	16 (1.4)	56 (1.8)	44 (1.8)
1996	New Jersey	2 (0.7)	20 (1.9)	64 (2.4)	36 (2.4)
	Northeast	2 (0.9)	23 (1.9)	67 (4.2)	33 (4.2)
	Nation	1 (0.4)	17 (1.2)	61 (1.7)	39 (1.7)

The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

Race/Ethnicity

Table 4.2 provides the percentages of fourth-grade public school students at or above each of the three mathematics achievement levels and also the percentage below the *Basic* level for White, Black, Hispanic, and Asian/Pacific Islander students.[40]

1996, Public School Students, Grade 4

The percentage of White students in New Jersey who attained the *Proficient* level was larger than that of Black and Hispanic students but was not significantly different from that of Asian/Pacific Islander students.

1992 vs. 1996, Public School Students, Grade 4

From 1992 to 1996, there was no significant change in the percentage of White, Black, Hispanic, or Asian/Pacific Islander students who attained the *Proficient* level in New Jersey.

[40] Results are reported for racial/ethnic subgroups meeting established sample size requirements (see Appendix A).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 4.2 — GRADE 4

Percentage of Public School Students Attaining Mathematics Achievement Levels by Race/Ethnicity

		Advanced	At or Above Proficient	At or Above Basic	Below Basic
White					
1992	New Jersey	3 (0.9)	32 (2.0)	81 (1.8)	19 (1.8)
	Northeast	3 (1.1)	30 (3.6)	75 (3.1)	25 (3.1)
	Nation	2 (0.4)	22 (1.5)	69 (1.4)	31 (1.4)
1996	New Jersey	4 (1.0)	36 (2.1)	84 (1.8)	16 (1.8)
	Northeast	4 (1.4)	33 (1.8)	81 (2.9)	19 (2.9)
	Nation	3 (0.5)	26 (1.3)	74 (1.6)	26 (1.6)
Black					
1992	New Jersey	0 (0.1)	3 (1.1)	29 (3.6)	71 (3.6)
	Northeast	0 (****)	4 (1.5)	26 (3.9)	74 (3.9)
	Nation	0 (****)	2 (0.7)	22 (1.9)	78 (1.9)
1996	New Jersey	0 (****)	3 (1.8)	35 (3.7)	65 (3.7)
	Northeast	0 (****)	6 (****)	29 (9.7)	71 (9.7)
	Nation	0 (0.1)	5 (1.5)	32 (3.4)	68 (3.4)
Hispanic					
1992	New Jersey	0 (****)	6 (2.0)	42 (4.4)	58 (4.4)
	Northeast	0 (****)	5 (1.5)	30 (5.3)	70 (5.3)
	Nation	0 (****)	5 (1.0)	33 (2.3)	67 (2.3)
1996	New Jersey	0 (****)	5 (2.0)	40 (4.6)	60 (4.6)
	Northeast	0 (****)	12 (3.8)	48 (6.3)	52 (6.3)
	Nation	0 (****)	7 (1.0)	40 (2.6)	60 (2.6)
Asian/Pacific Islander					
1992	New Jersey	6 (2.9)	40 (4.5)	83 (5.5)	17 (5.5)
	Northeast	*** (**.*)	*** (**.*)	*** (**.*)	*** (**.*)
	Nation	4 (2.0)	30 (4.9)	75 (3.5)	25 (3.5)
1996	New Jersey	8 (3.3)	48 (5.0)	92 (2.4)	8 (2.4)
	Northeast	*** (**.*)	*** (**.*)	*** (**.*)	*** (**.*)
	Nation	5 (2.8)	24 (6.0)	72 (5.5)	28 (5.5)

Results are reported for racial/ethnic subgroups meeting established sample size requirements (see Appendix A). The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. *** Sample size is insufficient to permit a reliable estimate. **** Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

Students' Reports of Parents' Highest Education Level

Table 4.3 shows the mathematics achievement level results for fourth-grade public school students who reported that neither parent graduated from high school, at least one parent graduated from high school, at least one parent had some education after high school, at least one parent graduated from college, or that they did not know their parents' highest education level. It should be noted that 32 percent of fourth graders in New Jersey reported that they did not know the education level of either of their parents.

1996, Public School Students, Grade 4

The percentage of New Jersey students who reported that neither parent graduated from high school who attained the *Proficient* level did not differ significantly from that of students who reported that at least one parent graduated from high school or at least one parent had some education after high school but was smaller than that of students who reported that at least one parent graduated from college.

1992 vs. 1996, Public School Students, Grade 4

The percentage of students in New Jersey who reported that at least one parent graduated from high school, at least one parent had some education after high school, or at least one parent graduated from college who attained the *Proficient* level did not change significantly between 1992 and 1996.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 4.3 — GRADE 4

Percentage of Public School Students Attaining Mathematics Achievement Levels by Students' Reports of Parents' Highest Education Level

	Advanced	At or Above Proficient	At or Above Basic	Below Basic
Did not finish high school				
1992 New Jersey	0 (****)	7 (****)	48 (7.7)	52 (7.7)
Northeast	*** (**.*)	*** (**.*)	*** (**.*)	*** (**.*)
Nation	0 (****)	5 (1.9)	38 (5.4)	62 (5.4)
1996 New Jersey	0 (****)	8 (5.2)	41 (8.1)	59 (8.1)
Northeast	*** (**.*)	*** (**.*)	*** (**.*)	*** (**.*)
Nation	0 (****)	5 (1.8)	36 (4.4)	64 (4.4)
Graduated from high school				
1992 New Jersey	1 (1.0)	14 (2.4)	60 (3.8)	40 (3.8)
Northeast	0 (****)	12 (4.9)	55 (7.5)	45 (7.5)
Nation	0 (****)	12 (1.9)	52 (2.9)	48 (2.9)
1996 New Jersey	1 (****)	18 (3.7)	65 (4.6)	35 (4.6)
Northeast	1 (****)	27 (4.7)	66 (7.3)	34 (7.3)
Nation	1 (****)	15 (2.0)	58 (3.0)	42 (3.0)
Some education after HS				
1992 New Jersey	2 (****)	27 (5.7)	75 (3.8)	25 (3.8)
Northeast	6 (2.9)	29 (4.7)	73 (7.2)	27 (7.2)
Nation	2 (0.8)	21 (2.5)	67 (3.0)	33 (3.0)
1996 New Jersey	1 (****)	25 (5.1)	81 (4.6)	19 (4.6)
Northeast	4 (****)	36 (6.2)	81 (4.9)	19 (4.9)
Nation	2 (1.1)	27 (3.1)	76 (3.0)	24 (3.0)
Graduated from college				
1992 New Jersey	4 (0.9)	34 (2.4)	77 (2.2)	23 (2.2)
Northeast	5 (1.8)	33 (4.6)	72 (3.4)	28 (3.4)
Nation	3 (0.6)	25 (2.0)	66 (1.5)	34 (1.5)
1996 New Jersey	4 (1.3)	34 (2.6)	74 (2.7)	26 (2.7)
Northeast	4 (1.4)	33 (3.3)	74 (4.2)	26 (4.2)
Nation	4 (0.7)	27 (1.9)	70 (1.9)	30 (1.9)
I don't know				
1992 New Jersey	1 (****)	16 (2.0)	58 (3.1)	42 (3.1)
Northeast	0 (****)	15 (2.3)	54 (3.4)	46 (3.4)
Nation	1 (0.3)	12 (1.1)	50 (1.5)	50 (1.5)
1996 New Jersey	1 (0.6)	16 (1.6)	60 (3.1)	40 (3.1)
Northeast	2 (****)	17 (2.8)	62 (4.8)	38 (4.8)
Nation	1 (0.4)	14 (1.2)	55 (2.1)	45 (2.1)

Results are reported for parental education subgroups meeting established sample size requirements (see Appendix A). The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. *** Sample size is insufficient to permit a reliable estimate. **** Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

Title I Participation

Table 4.4 presents the percentage of fourth graders at or above each of the mathematics achievement levels as well as the percentage of students below *Basic* by Title I participation. (Results based on participation in Title I programs are not available for previous state-level mathematics assessments.)

1996, Public School Students, Grade 4

In New Jersey 2 percent of the students receiving Title I services performed at or above the *Proficient* level. Less than one third of the students who did not receive Title I services attained the *Proficient* level (29 percent).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 4.4 — GRADE 4

Percentage of Public School Students Attaining Mathematics Achievement Levels by Title I Participation

		Advanced	At or Above Proficient	At or Above Basic	Below Basic
GRADE 4					
Participated					
1996	New Jersey	0 (****)	2 (1.4)	23 (3.9)	77 (3.9)
	Northeast	0 (****)	3 (1.3)	29 (5.2)	71 (5.2)
	Nation	0 (****)	3 (0.9)	31 (2.7)	69 (2.7)
Did not participate					
1996	New Jersey	3 (0.8)	29 (1.9)	75 (2.0)	25 (2.0)
	Northeast	3 (1.3)	31 (2.1)	77 (3.4)	23 (3.4)
	Nation	3 (0.4)	25 (1.3)	72 (1.6)	28 (1.6)

Results are reported for students participating in Title I programs only if established sample size requirements are met (see Appendix A). The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). **** Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

Free/Reduced-Price Lunch Program Eligibility

Table 4.5 shows 1996 mathematics achievement level results for fourth graders based on their eligibility for the federally funded free/reduced-price lunch component of the National School Lunch Program. (Similar results are not available for previous NAEP mathematics assessments.)

1996, Public School Students, Grade 4

The percentage of students in New Jersey eligible for free or reduced-price lunch who performed at or above the *Proficient* level (5 percent) was not significantly different from that of students nationwide (8 percent). The percentage of students in New Jersey who were not eligible for this service who attained this level (35 percent) was greater than the figure for the nation (25 percent). In New Jersey, the percentage of students eligible for free or reduced-price lunch who attained the *Proficient* level was smaller than that of students who were not.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 4.5 — GRADE 4

Percentage of Public School Students Attaining Mathematics Achievement Levels by Free/Reduced-Price Lunch Eligibility

		Advanced	At or Above Proficient	At or Above Basic	Below Basic
GRADE 4					
Eligible					
1996	New Jersey	0 (****)	5 (1.5)	40 (3.3)	60 (3.3)
	Northeast	1 (****)	10 (2.8)	42 (7.4)	58 (7.4)
	Nation	0 (0.3)	8 (1.2)	41 (2.6)	59 (2.6)
Not eligible					
1996	New Jersey	4 (1.0)	35 (2.1)	81 (2.0)	19 (2.0)
	Northeast	4 (1.4)	33 (2.4)	82 (2.7)	18 (2.7)
	Nation	3 (0.6)	25 (1.4)	73 (1.8)	27 (1.8)

The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). *** Sample size is insufficient to permit a reliable estimate. **** Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

Type of Location

Table 4.6 presents achievement level results for fourth-grade students attending public schools in central cities and urban fringes/large towns.[41]

1996, Public School Students, Grade 4

The percentage of New Jersey students attending schools in central cities who attained the *Proficient* level was smaller than that of students in urban fringes/large towns.

1992 vs. 1996, Public School Students, Grade 4

From 1992 to 1996, there was no significant change in the percentage of students attending schools in central cities or urban fringes/large towns in New Jersey who attained the *Proficient* level.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 4.6 — GRADE 4

Percentage of Public School Students Attaining Mathematics Achievement Levels by Type of Location

		Advanced	At or Above Proficient	At or Above Basic	Below Basic
Central city					
1992	New Jersey	0 (****)	7 (2.0)	34 (5.5)	66 (5.5)
	Nation	1 (0.3)	13 (1.4)	48 (2.3)	52 (2.3)
1996	New Jersey	1 (****)	14 (4.6)	49 (8.5)	51 (8.5)
	Nation	2 (0.6)	15 (1.4)	51 (3.7)	49 (3.7)
Urban fringe/Large town					
1992	New Jersey	3 (0.7)	28 (1.9)	74 (2.2)	26 (2.2)
	Nation	3 (0.6)	23 (2.1)	66 (2.1)	34 (2.1)
1996	New Jersey	3 (0.8)	26 (2.1)	71 (2.7)	29 (2.7)
	Nation	3 (0.6)	24 (1.8)	69 (2.3)	31 (2.3)

Results are reported for type of location subgroups meeting established sample size requirements (see Appendix A). The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. Characteristics of the school sample do not permit reliable regional results for type of location. **** Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

[41] Results are reported for type of location subgroups meeting established sample size requirements (see Appendix A).

PART THREE

Finding a Context for Understanding Students' Mathematics Performance in Public Schools

The mathematics performance of public school students in New Jersey often can be better understood when viewed in the context of the environment in which the students are learning. This educational environment is largely determined by school characteristics, by characteristics of mathematics instruction in the school, by home support for academics and other home influences, and by the students' own views about mathematics. Information about this environment is gathered by means of the questionnaires administered to principals, teachers, and students.

Because NAEP is administered to a sample of students that is representative of the fourth-grade student populations in the schools of New Jersey, NAEP results provide a broad view of the educational practices in New Jersey, which is useful for improving instruction and setting policy. However, despite the richness of the NAEP results, it is very important to note that NAEP data cannot establish a cause-and-effect relationship between educational environment and student scores on the NAEP mathematics assessment.

The variables contained in Part Three are from the school characteristics and policies questionnaire, teacher questionnaires, and student background questions. These questionnaires are sometimes refined by revising or reformatting questions between assessments. Additional questions are sometimes added to questionnaires to track emerging trends in education. These revisions and additions may make comparisons with results from previous years unwise or impossible. When appropriate, trends across assessments are presented; otherwise, only 1996 results are discussed.

Part Three consists of three chapters: Chapter 5 discusses school characteristics related to mathematics instruction;[42] Chapter 6 describes classroom practices related to mathematics instruction, including calculator and computer use; and Chapter 7 covers some potential influences from the home and from the students' own views about mathematics.

[42] Information on teacher preparation is included in Appendix E of this report.

CHAPTER 5

School Characteristics Related to Mathematics Instruction

School programs and conditions, instructional practices, and resource availability vary from state to state and even among schools within a locality. The information in this chapter is intended to give insight into those characteristics that are associated with students' success in mathematics.

The variables reported here reflect information from the questionnaires completed by principals and teachers of the public school students in the NAEP 1996 mathematics assessment. In all cases, analyses are done at the student level. School and teacher-reported results are given in terms of the percentage of students who attend schools or who have teachers reporting particular practices.[43]

Emphasis on Mathematics in the School

In the school characteristics and policies questionnaire, principals were asked whether their school has identified mathematics as a priority in the last two years (i.e., whether mathematics receives special emphasis in school-wide goals and objectives, instruction, and workshops, etc.). Table 5.1 presents the public school principals' reports.

- The percentage of fourth-grade students in New Jersey who attended schools that reported that mathematics was a priority (81 percent) was not significantly different from the national percentage (76 percent).

[43] Appendix A provides more details on the units of analysis used to derive the results presented in this report.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 5.1 — GRADE 4

Public Schools' Reports on Mathematics as a Priority

Has your school identified mathematics as a priority in the last two years?	**New Jersey**	**Northeast**	**Nation**
	Percentage and Average Scale Score		
Yes	81 (5.1) 227 (2.2)	86 (3.9) 228 (2.9)	76 (3.9) 222 (1.2)
No	19 (5.1) 230 (2.7)!	14 (3.9) 225 (5.1)!	24 (3.9) 220 (3.1)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

Resource Availability to Teachers

Resources available to teachers and schools vary. Past surveys have shown that teachers' perceptions of the availability of resources (e.g., instructional materials, staff, and preparation or planning time) are variable across the country.[44] Previous NAEP assessments in several subject areas have also shown a positive relationship between teachers' reports of resource availability and their students' performance in most states.[45]

Availability of Instructional Materials

Teachers often see the lack of resources as a key problem for mathematics instruction. In 1993 a national survey reported that the average school spent $100 per year on mathematics software, $1.00 per elementary school student on mathematics materials and manipulatives, and $0.40 per middle school student on mathematics materials.[46] Teachers were asked to categorize how well their school systems provided them with the classroom instructional materials they needed. Table 5.2 shows the percentages of fourth-grade students whose teachers reported receiving varying levels of support from their public schools.

- The average mathematics scale score of students in New Jersey whose teachers reported receiving all the resources they needed (231) was greater than that of students whose teachers received some or none of the resources they needed (219). The percentage of students whose teachers reported receiving all of the resources they needed in New Jersey (17 percent) was not significantly different from that of students across the nation (12 percent).

[44] National Center for Education Statistics. *Schools and Staffing in the United States: A Statistical Profile, 1993-94*. (Washington, DC: National Center for Education Statistics, 1996).

[45] Miller, K.E., J.E. Nelson, and M. Naifeh. *Cross-State Data Compendium for the NAEP 1994 Grade 4 Reading Assessment*. (Washington, DC: National Center for Education Statistics, 1995).

[46] Council of Chief State School Officers. *State Indicators of Science and Mathematics Education 1995*. (Washington, DC: Council of Chief State School Officers, State Education Assessment Center, 1995).

TABLE 5.2 — GRADE 4

Public School Teachers' Reports on Resource Availability

Which of the following statements is true about how well your school system provides you with the instructional materials and other resources you need to teach your class?		New Jersey	Northeast	Nation
		Percentage and Average Scale Score		
I get some or none of the resources I need.	1992	38 (3.4) 220 (2.9)	44 (6.1) 216 (4.2)	37 (3.5) 214 (2.0)
	1996	27 (3.9) 219 (3.2)	36 (5.7) 219 (5.0)!	34 (2.5) 221 (1.5)>
I get most of the resources I need.	1992	47 (3.7) 232 (1.5)	45 (5.4) 229 (3.3)	52 (3.0) 221 (1.2)
	1996	57 (3.3) 231 (2.0)	46 (2.7) 233 (3.1)	55 (2.2) 224 (1.5)
I get all the resources I need.	1992	15 (3.0) 231 (3.7)!	11 (3.3) 234 (3.8)!	11 (1.7) 222 (2.7)
	1996	17 (3.2) 231 (4.0)	18 (4.8) 230 (3.5)!	12 (1.8) 224 (2.3)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

Availability of Curriculum Specialist in the School

Teachers were asked if there was a curriculum specialist available to help or advise them in mathematics. Table 5.3 shows the public school fourth-grade teachers' responses.

- In New Jersey, more than half of the students were taught by teachers who reported that there was a curriculum specialist available to help or advise in mathematics (58 percent). This figure was greater than that of students across the nation (46 percent).
- The percentage of New Jersey students whose teachers reported having a curriculum specialist available in mathematics did not change significantly from 1992 (59 percent) to 1996 (58 percent).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 5.3 — GRADE 4

Public School Teachers' Reports on Curriculum Specialists

Is there a curriculum specialist available to help or advise you in mathematics?		New Jersey	Northeast	Nation
		Percentage and Average Scale Score		
Yes	1992	59 (4.0) 224 (2.2)	45 (8.3) 223 (2.9)!	50 (3.3) 218 (1.3)
	1996	58 (4.3) 226 (2.1)	53 (9.0) 225 (5.9)!	46 (3.6) 221 (2.1)
No	1992	41 (4.0) 231 (1.9)	55 (8.3) 223 (4.1)	50 (3.3) 219 (1.6)
	1996	42 (4.3) 230 (2.4)	47 (9.0) 228 (4.0)!	54 (3.6) 224 (1.4)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

In-School Teacher Preparation Time

Teachers were asked to indicate how many school hours they had designated as preparation time per week. The question did not restrict preparation time to the mathematics classes of those students who took part in the NAEP assessment or mathematics classes in general. The question referred to the preparation time allotted for all classes taught by the teacher. The responses of public school fourth-grade teachers are shown in Table 5.4.

- The percentage of students in New Jersey whose teachers reported having five or more school hours designated as preparation time per week (17 percent) was smaller than that of students across the nation (27 percent).

- The percentage of New Jersey students whose teachers had one to two hours to prepare each week (22 percent) was not significantly different from that of students nationwide (21 percent).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 5.4 — GRADE 4

Public School Teachers' Reports on Preparation Time

How many school hours do you have designated as preparation time per week?	New Jersey	Northeast	Nation
	Percentage and Average Scale Score		
Less than 1	1 (0.6) *** (**.*)	1 (****) *** (**.*)	7 (1.4) 219 (4.2)!
1-2	22 (3.3) 228 (2.7)	16 (4.9) 228 (3.5)!	21 (2.1) 222 (2.4)
3-4	59 (3.7) 227 (1.9)	50 (6.0) 229 (3.1)	45 (2.5) 226 (1.4)
5 hours or more	17 (2.6) 228 (3.7)	32 (5.6) 226 (3.7)	27 (2.6) 221 (2.1)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic. *** Sample size is insufficient to permit a reliable estimate. **** Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

Parents as Classroom Aides

When school personnel and parents develop a positive line of communication, they strengthen the learning environment for the students both at school and at home. One of the most frequent reasons cited by school personnel for contacting parents is to request parent volunteer time at school.[47] The principals of the participating public schools were asked if parents were used as classroom aides. As shown in Table 5.5, principals for fourth graders reported the following.

- About one fifth of the students in New Jersey (21 percent) were in schools that reported routinely using parents as aides in classrooms. In contrast, 28 percent of students in New Jersey attended schools where parents were not used as classroom aides.

- From 1992 to 1996, the percentage of New Jersey students in schools that reported routinely using parents as aides in classrooms did not change significantly* (15 percent in 1992 and 21 percent in 1996).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 5.5 — GRADE 4

Public Schools' Reports on Parents as Aides in Classrooms

Does your school use parents as aides in classrooms?		**New Jersey**	**Northeast**	**Nation**
		Percentage and Average Scale Score		
No	1992	49 (4.8) 226 (1.9)	12 (5.8) 222 (7.4)!	11 (2.5) 212 (4.5)!
	1996	28 (5.0)< 224 (4.4)	10 (3.4) 231 (5.9)!	8 (1.7) 219 (3.7)!
Yes, occasionally	1992	37 (4.5) 227 (3.1)	42 (8.0) 222 (4.3)	42 (3.3) 216 (1.4)
	1996	50 (5.8) 228 (2.6)	32 (10.1) 221 (2.4)!	44 (4.8) 219 (2.1)
Yes, routinely	1992	15 (3.7) 234 (4.3)!	46 (8.8) 226 (2.6)!	47 (3.4) 223 (1.5)
	1996	21 (4.7) 231 (4.6)!	58 (10.5) 231 (5.1)	48 (4.9) 225 (1.9)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

* Although the difference may appear large, recall that "significance" here refers to "statistical significance."

[47] National Center for Education Statistics. *The Condition of Education 1995*. (Washington, DC: National Center for Education Statistics, 1995).

Student Absenteeism

School principals were asked if student absenteeism was a serious, moderate, or minor problem, or not a problem. Table 5.6 shows, for fourth graders, results based on principals' reports.

- In New Jersey, 9 percent of the students attended public schools that reported that absenteeism was a moderate to serious problem. This percentage was not significantly different from that of students across the nation (13 percent).
- The percentage of students in schools reporting that absenteeism was a moderate to serious problem in New Jersey did not change significantly from 1992 (8 percent) to 1996 (9 percent).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 5.6 — GRADE 4

Public Schools' Reports on Student Absenteeism

To what degree is student absenteeism a problem in your school?		New Jersey	Northeast	Nation
		Percentage and Average Scale Score		
Not a problem	1992	59 (5.1) 232 (2.0)	42 (6.2) 232 (3.5)	31 (2.7) 226 (1.7)
	1996	66 (5.4) 234 (2.2)	62 (9.6) 234 (2.7)	43 (4.1) 230 (1.8)
Minor	1992	33 (4.8) 224 (3.1)	43 (7.6) 221 (3.7)!	54 (3.1) 219 (1.4)
	1996	26 (4.6) 218 (3.6)	38 (9.6) 216 (7.1)!	44 (4.3) 218 (2.3)
Moderate to serious	1992	8 (2.8) 203 (6.7)!	14 (5.4) 206 (6.8)!	15 (2.2) 202 (2.1)
	1996	9 (3.3) 201 (7.2)!	0 (****) *** (**.*)	13 (2.3) 210 (3.1)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic. *** Sample size is insufficient to permit a reliable estimate. **** Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

CHAPTER 6

Classroom Practices Related to Mathematics Instruction

The mathematics achievement of our nation's students has been the topic of considerable discussion in recent years. The mathematics achievement of our students does not compare well with that of students in other countries[48] or with achievement goals set by the United States for itself.[49] Improvements in mathematics performance during the 1980s and 1990s are encouraging;[50] however, policy makers and educators must continue to evaluate the state of mathematics education and to commit to improvements to school mathematics programs.

For some of the issues discussed in this chapter, student- and teacher-reported results for similar questions are presented. In these situations, some discrepancies may exist between student- and teacher-reported percentages. It is not possible to offer conclusive reasons for these discrepancies or to determine which reports most accurately reflect fourth-grade classroom activities. The reports presented represent students' and teachers' impressions of the frequency of various activities in the classrooms.

An important step in the improvement of mathematics education in the nation's elementary and secondary schools was the development and adoption of the National Council of Teachers of Mathematics (NCTM) *Curriculum and Evaluation Standards for School Mathematics*.[51] Already adopted by the majority of the states,[52] the NCTM *Standards* represent a basis upon which mathematics instruction can be reformed and improved. This chapter focuses on curricular and instructional content issues in New Jersey public schools and their relationship to students' mathematics performance.

---

[48] Lapointe, A.E., N.A. Mead, and J.M. Askew. *Learning Mathematics*. (Washington, DC: The International Assessment of Educational Progress, National Center for Education Statistics, 1992); Beaton, A.E., I.V.S. Mullis, M.D. Martin, E.S. Gonzalez, D.L. Kelly, and T.A. Smith. *Mathematics Achievement in the Middle School Years*. (Chestnut Hill, MA: TIMSS International Study Center, Boston College, 1996).

[49] Reese, C.M., K.E. Miller, J. Mazzeo, and J.A. Dossey. *NAEP 1996 Mathematics Report Card*. (Washington, DC: National Center for Education Statistics, 1997).

[50] Campbell, J.R., C.M. Reese, C. O'Sullivan, and J.A. Dossey. *NAEP 1994 Trends in Academic Progress*. (Washington, DC: National Center for Education Statistics, 1996).

[51] National Council of Teachers of Mathematics. *Curriculum and Evaluation Standards for School Mathematics*. (Reston, VA: NCTM, 1989).

[52] Council of Chief State School Officers. *Analysis of State Education Indicators: State Profiles and NAEP Results Related to State Policies and Practices. (Washington, DC: Council of Chief State School Officers, State Education Assessment Center, 1993).*

NCTM Standards

Since their publication in 1989, the NCTM *Standards* have received considerable attention by policy makers and educators. The NCTM *Standards* outline curriculum and evaluation recommendations for kindergarten through grade 12 mathematics instruction. To gauge how knowledgeable teachers are about the standards, teachers were asked about their familiarity with the NCTM *Standards* and their involvement in professional development related to them. Table 6.1 shows the results based on the responses of fourth-grade public school mathematics teachers.

- A small percentage of the fourth-grade students in New Jersey (5 percent) had mathematics teachers who reported being very knowledgeable about the NCTM *Standards*. This percentage was smaller than the percentage whose teachers reported having little or no knowledge of the *Standards* (42 percent).

- The percentage of New Jersey fourth graders whose teachers reported being very knowledgeable about the NCTM *Standards* did not differ significantly from the percentage for the nation (6 percent).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 6.1 — GRADE 4

Public School Teachers' Reports on Knowledge of the NCTM Standards

How knowledgeable are you about the National Council of Teachers of Mathematics (NCTM) Curriculum and Evaluation Standards for School Mathematics?	New Jersey	Northeast	Nation
	Percentage and Average Scale Score		
I have little or no knowledge.	42 (3.7) 225 (1.9)	48 (4.6) 226 (3.2)	45 (2.4) 222 (1.7)
Somewhat knowledgeable	38 (3.8) 227 (2.8)	28 (4.8) 219 (4.0)!	32 (2.4) 222 (1.6)
Knowledgeable	14 (2.4) 237 (2.8)	15 (3.6) 229 (3.1)!	18 (2.0) 222 (2.0)
Very knowledgeable	5 (1.9) 232 (4.9)!	9 (3.2) 249 (10.1)!	6 (1.2) 235 (4.8)!

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

Teachers were also asked whether they had participated in any professional development activities that provided them with strategies for implementing the NCTM *Standards*. The activities could include local workshops, and regional and national NCTM meetings. (Teachers were asked to select all activities that applied.) Table 6.2 presents the results for fourth-grade public school mathematics teachers.

- More than half of the students in New Jersey (54 percent) had mathematics teachers who reported attending no professional development activities related to implementing the NCTM *Standards*. This percentage did not differ significantly from the percentage for the nation (60 percent).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 6.2 — GRADE 4

Public School Teachers' Reports on Professional Development for Implementing the NCTM Standards

Have you participated in any professional development activities that have provided you with strategies for implementing the 1989 NCTM Curriculum and Evaluation Standards for School Mathematics?	New Jersey	Northeast	Nation
	Percentage and Average Scale Score		
Local workshop	34 (4.3) 231 (3.0)	23 (5.4) 228 (6.4)!	29 (2.5) 223 (1.9)
Regional or national NCTM meeting	9 (2.7) 232 (3.2)!	8 (2.1) *** (**.*)	9 (1.4) 229 (3.1)
Other	11 (2.5) 231 (4.1)!	20 (4.4) 236 (7.3)!	12 (1.7) 230 (3.3)
No	54 (4.4) 224 (1.8)	61 (5.1) 225 (2.0)	60 (2.8) 222 (1.3)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic. *** Sample size is insufficient to permit a reliable estimate.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

Course-Taking Patterns

To investigate the relationship between the mathematics performance of students and their study of mathematics in school, information on the amount of time each week spent on mathematics instruction in class was collected.

The amount of time spent on mathematics instruction within the classroom varies from school to school and from state to state.[53] The teachers of the fourth graders participating in the NAEP assessment were asked about the amount of time spent each week on mathematics instruction. Table 6.3 presents the public school teachers' responses.

- In 1996, about two thirds of the students in New Jersey had teachers who reported spending four or more hours on mathematics instruction (71 percent), compared to 3 percent who reported spending two and a half hours or less. The average scale score for students receiving four or more hours of mathematics instruction (227) was not significantly different from* that for students receiving two and a half hours or less (214).

- The percentage of fourth graders in New Jersey whose teachers reported spending four hours or more on mathematics instruction was not significantly different from the percentage for the nation (69 percent).

- The percentage of New Jersey fourth graders in 1996 whose teachers reported spending four hours or more on mathematics instruction did not differ significantly from the percentage in 1992 (69 percent).

---

* Although the difference may appear large, recall that "significance" here refers to "statistical significance."

[53] National Center for Education Statistics. *Schools and Staffing in the United States: A Statistical Profile, 1993-94*. (Washington, DC: National Center for Education Statistics, 1996).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 6.3 — GRADE 4

Public School Teachers' Reports on Time Spent on Mathematics Instruction

How much time do you spend each week on mathematics instruction with this class?		New Jersey	Northeast	Nation
		Percentage and Average Scale Score		
2½ hours or less	1992	3 (1.4) *** (**.*)	6 (2.8) *** (**.*)	5 (1.0) 224 (3.7)!
	1996	3 (1.3) 214 (8.8)!	3 (****) *** (**.*)	5 (1.1) 225 (2.8)!
More than 2½ hours but less than 4 hours	1992	29 (4.4) 231 (2.3)	21 (5.6) 219 (3.7)!	21 (2.1) 222 (2.6)
	1996	26 (4.2) 230 (1.8)	27 (5.9) 229 (2.8)!	25 (2.5) 225 (1.9)
4 hours or more	1992	69 (4.6) 227 (1.9)	73 (6.8) 222 (3.6)	74 (2.5) 217 (1.1)
	1996	71 (3.9) 227 (2.0)	70 (7.4) 224 (3.4)	69 (2.8) 222 (1.1)>

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic. *** Sample size is insufficient to permit a reliable estimate. **** Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

Instructional Emphasis

The framework or blueprint that guided the development of the NAEP 1996 mathematics assessment identified three mathematical abilities — conceptual understanding, procedural knowledge, and problem solving. The NCTM *Standards* emphasize the need to give students the opportunity to develop all three abilities. Focusing on only one of these abilities is limiting to students' mathematical development. For example, emphasizing how to do a problem (procedural knowledge) without understanding why it works (conceptual understanding) or when to use it (problem solving) offers students, at best, an incomplete picture of mathematics. These three abilities were reflected in the specifications for the 1990, 1992, and 1996 NAEP mathematics assessments. Questions assessing all of these abilities were included. Also identified by the NCTM *Standards* is the ability to reason which depends upon all three of the previous abilities, and the need for students to be able to communicate mathematical ideas. To assess the students' opportunities to experience a variety of learning, teachers were asked how often they addressed skills related to these abilities. Table 6.4 shows the results based on responses given by grade 4 teachers.

- Almost all of the fourth graders had teachers who reported they addressed the learning of mathematics facts and concepts a lot (94 percent), while none (0 percent) had teachers who reported spending little or no time on the topic.

- Almost all of the students had teachers who reported they addressed the learning of skills and procedures a lot (93 percent). At the other extreme, 1 percent of the students had teachers who reported spending little or no time on the topic.

- Teachers of 59 percent of the students reported they addressed the developing of reasoning and analytical ability a lot. In contrast, 5 percent had teachers who reported spending little or no time addressing the topic.

- In terms of addressing the learning of how to communicate ideas in mathematics clearly, 46 percent of fourth graders had teachers who reported doing so a lot, while 15 percent of the students had teachers who reported spending little or no time addressing the topic.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 6.4 — GRADE 4

Public School Teachers' Reports on Skills Addressed

In this mathematics class, how often do you address each of the following?	New Jersey	Northeast	Nation
	Percentage and Average Scale Score		
Learning mathematics facts and concepts			
A little or none	0 (0.3) *** (**.*)	1 (****) *** (**.*)	0 (****) *** (**.*)
Some	5 (1.4) 221 (6.6)!	8 (2.1) *** (**.*)	7 (1.1) 221 (3.5)
A lot	94 (1.4) 228 (1.5)	91 (2.0) 226 (2.6)	93 (1.1) 223 (1.1)
Learning skills and procedures needed to solve routine problems			
A little or none	1 (0.5) *** (**.*)	0 (****) *** (**.*)	0 (****) *** (**.*)
Some	5 (1.3) 219 (4.2)!	6 (3.0) *** (**.*)	9 (1.3) 220 (3.9)
A lot	93 (1.4) 228 (1.6)	94 (3.0) 227 (2.4)	91 (1.3) 223 (1.1)
Developing reasoning and analytical ability to solve unique problems			
A little or none	5 (1.3) 221 (4.9)!	13 (4.3) 227 (3.4)!	7 (1.2) 222 (3.4)
Some	36 (4.4) 225 (2.3)	35 (3.7) 218 (4.6)	39 (2.6) 219 (1.7)
A lot	59 (4.2) 229 (2.2)	52 (3.7) 232 (2.8)	54 (2.5) 226 (1.6)
Learning how to communicate ideas in mathematics effectively			
A little or none	15 (2.8) 229 (3.1)	17 (4.3) 229 (2.8)!	17 (1.9) 227 (2.1)
Some	39 (3.8) 226 (2.8)	35 (6.4) 229 (5.0)!	44 (2.7) 219 (2.0)
A lot	46 (3.5) 229 (2.1)	47 (7.6) 223 (2.7)	39 (2.6) 225 (1.6)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic. *** Sample size is insufficient to permit a reliable estimate. **** Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

Communicating Mathematical Ideas

Much focus in the mathematics education reform effort has been placed on students' ability to communicate their understanding of mathematics to others. Results presented previously in Table 6.4 examine the level of emphasis teachers placed on communication in their classroom. As a follow-up, the students participating in the NAEP assessment, and their teachers, were asked about how often they were asked to write a few sentences about solving a mathematical problem and how often they were asked to discuss solutions to mathematics problems with other students.

Based on the responses of the fourth-grade public school teachers in New Jersey, the results are shown in Table 6.5.

- A small percentage of the students were asked to write about solving a mathematics problem (6 percent) and less than half were asked to discuss solutions with other students (42 percent) almost every day. By comparison, 23 percent were never or hardly ever asked to write about their solutions and 4 percent were never or hardly ever asked to discuss their solutions.

- The average scale score for students who were asked to discuss solving a mathematics problem almost every day (230) was not significantly different from that for students who never or hardly ever discussed their solutions (225).

When New Jersey fourth-grade public school students were asked about how often they wrote about or discussed solutions to mathematics problems, they reported the following:

- About one fifth of the students said they were asked to write about mathematics solutions (20 percent) and about one fifth said they were asked to discuss their solutions with other students (22 percent) almost every day. At the other end of the continuum, 33 percent of students said they were never or hardly ever asked to write about solutions and 29 percent said they were never or hardly ever asked to discuss their solutions.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 6.5 — GRADE 4

Public School Teachers' and Students' Reports on Writing About and Discussing Mathematics Problems

How often do the students in this class (do you) do each of the following?	New Jersey		Northeast		Nation	
	Teacher	Student	Teacher	Student	Teacher	Student
	Percentage and Average Scale Score					
Write a few sentences about how to solve a mathematics problem						
Never or hardly ever	23 (3.2) 226 (3.0)	33 (1.6) 230 (1.8)	22 (4.5) 227 (4.5)!	34 (3.3) 231 (2.8)	27 (2.4) 220 (1.7)	33 (1.4) 226 (1.3)
Once or twice a month	44 (3.7) 227 (2.4)	18 (0.8) 233 (2.2)	29 (6.9) 222 (11.5)!	15 (1.5) 235 (3.1)	35 (2.8) 223 (2.5)	17 (0.7) 231 (1.3)
Once or twice a week	27 (3.2) 229 (2.7)	29 (1.3) 230 (2.0)	32 (7.2) 224 (3.9)!	31 (2.8) 229 (2.9)	27 (2.2) 223 (1.8)	27 (0.9) 223 (1.6)
Almost every day	6 (1.7) 231 (6.5)!	20 (1.4) 214 (2.3)	18 (4.4) 234 (6.1)!	21 (2.3) 212 (4.1)	10 (1.6) 233 (3.2)	22 (1.2) 211 (1.9)
Discuss solutions to mathematics problems with other students						
Never or hardly ever	4 (1.5) 225 (3.7)!	29 (1.6) 227 (2.0)	4 (2.5) *** (**.*)	32 (2.6) 226 (2.8)	6 (1.4) 219 (3.0)!	33 (1.0) 222 (1.0)
Once or twice a month	14 (2.6) 221 (5.2)	17 (1.0) 232 (2.1)	25 (4.5) 220 (6.0)	18 (1.3) 230 (3.7)	22 (1.7) 221 (1.9)	18 (0.6) 227 (1.7)
Once or twice a week	39 (4.0) 228 (2.2)	32 (1.3) 230 (2.1)	34 (3.8) 225 (4.1)	28 (1.6) 231 (2.9)	37 (2.3) 221 (1.9)	29 (0.7) 224 (1.4)
Almost every day	42 (3.9) 230 (2.0)	22 (1.1) 222 (2.0)	37 (3.9) 230 (4.9)	22 (2.0) 220 (3.2)	35 (2.1) 227 (1.8)	19 (0.8) 217 (1.5)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic. *** Sample size is insufficient to permit a reliable estimate.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

Collaboration in Small Groups

In many subject areas, researchers have found benefits from having students work collaboratively in small groups.[54] To examine the extent to which small groups are being used in instruction, students and their mathematics teachers were asked about the prevalence of these practices.

Table 6.6 shows the following based on the reports of fourth-grade public school mathematics teachers:

- A large majority of the fourth-grade students in New Jersey worked mathematics problems in small groups every day (28 percent) or once or twice a week (54 percent); a small percentage of the fourth graders never or hardly ever worked in groups (2 percent).

According to fourth graders' responses:

- In New Jersey, 10 percent worked mathematics problems in small groups every day and another 29 percent worked in small groups once or twice a week. Less than half of the fourth graders reported never or hardly ever working in groups (37 percent).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 6.6 — GRADE 4

Public School Teachers' and Students' Reports on Solving Mathematics Problems in Small Groups or With a Partner

How often do the students in this class (do you) solve mathematics problems in small groups or with a partner?	New Jersey		Northeast		Nation	
	Teacher	Student	Teacher	Student	Teacher	Student
	Percentage and Average Scale Score					
Never or hardly ever	2 (1.0) *** (**.*)	37 (1.9) 226 (1.8)	9 (4.5) *** (**.*)	44 (1.4) 226 (2.8)	7 (1.6) 224 (2.9)!	42 (1.3) 222 (1.0)
Once or twice a month	17 (2.7) 231 (3.4)	24 (1.3) 237 (1.8)	7 (1.3) *** (**.*)	24 (2.0) 239 (2.3)	18 (1.5) 218 (1.7)	22 (0.7) 233 (1.3)
Once or twice a week	54 (3.8) 228 (2.0)	29 (1.4) 227 (1.4)	59 (5.6) 221 (2.8)	24 (1.3) 226 (1.5)	50 (2.2) 222 (1.6)	26 (1.0) 221 (1.4)
Almost every day	28 (3.4) 226 (2.2)	10 (1.1) 210 (3.3)	25 (3.7) 236 (4.2)!	8 (1.3) 194 (3.9)	25 (2.1) 227 (1.9)	10 (0.7) 205 (2.3)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic. *** Sample size is insufficient to permit a reliable estimate.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

[54] Mulryan, C.M. "Fifth and Sixth Graders' Involvement and Participation in Cooperative Small Groups in Mathematics," in *Elementary School Journal, 95, 4*. (1995). pp. 297-310; Reuman, D.A. and D.J. MacIver. *Effects on Instructional Grouping on Seventh Graders' Academic Motivation and Achievement*. (Washington, DC: Office of Educational Research and Improvement, 1994).

Mathematics Homework

To examine the relationship between homework and mathematics performance, teachers of assessed students were asked to report the amount of mathematics homework they assigned each day, and students were asked to report the amount of time they spent on mathematics homework each day.

Table 6.7 shows the teachers' and students' responses for fourth-grade public school students in New Jersey. (Students had an additional response choice "I am not taking mathematics this year," but no analogous option was available to teachers.) According to fourth-grade teachers' responses:

- In New Jersey, almost all of the fourth graders in 1996 were assigned 15 minutes (38 percent) or 30 minutes (56 percent) of mathematics homework each day.

- None (0 percent) were not assigned any mathematics homework each day.

According to students in the fourth grade:

- A small percentage of the fourth graders did not spend any time on mathematics homework on a typical day (2 percent). By comparison, 49 percent spent 15 minutes and 26 percent spent 30 minutes on their mathematics homework.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 6.7 — GRADE 4

Public School Teachers' and Students' Reports on Homework

Approximately how much mathematics homework do you assign (time do you spend on math homework) each day?		New Jersey		Northeast		Nation	
		Teacher	Student	Teacher	Student	Teacher	Student
		Percentage and Average Scale Score					
I am not taking mathematics this year.	1992	--- (--.-) --- (--.-)	0 (0.1) *** (**.*)	--- (--.-) --- (--.-)	0 (0.2) *** (**.*)	--- (--.-) --- (--.-)	1 (0.1) *** (**.*)
	1996	--- (--.-) --- (--.-)	0 (0.1) *** (**.*)	--- (--.-) --- (--.-)	1 (0.4) *** (**.*)	--- (--.-) --- (--.-)	1 (0.2) *** (**.*)
None	1992	0 (****) *** (**.*)	2 (0.3) *** (**.*)	2 (****) *** (**.*)	3 (0.5) *** (**.*)	6 (1.4) 221 (2.6)!	7 (0.6) 222 (2.3)
	1996	0 (****) *** (**.*)	2 (0.5) *** (**.*)	1 (****) *** (**.*)	3 (0.8) *** (**.*)	4 (0.8) 232 (3.8)	6 (0.6) 226 (3.2)
15 minutes	1992	44 (4.2) 232 (2.1)	49 (1.4) 231 (1.6)	50 (6.2) 225 (4.6)	44 (3.4) 227 (3.0)	53 (2.1) 221 (1.4)	39 (1.1) 221 (1.2)
	1996	38 (3.9) 232 (2.0)	49 (1.6) 232 (1.5)	43 (5.1) 232 (5.4)	42 (2.5) 229 (2.6)	50 (2.6) 225 (1.6)	40 (1.1) 224 (1.3)
30 minutes	1992	52 (4.3) 225 (2.5)	27 (1.0) 229 (1.9)	44 (5.7) 221 (3.0)	29 (2.7) 230 (3.0)	36 (2.6) 216 (1.8)	29 (0.8) 222 (1.0)
	1996	56 (4.1) 227 (2.2)	26 (1.3) 229 (1.6)	51 (5.2) 221 (3.8)	31 (1.9) 234 (2.5)	40 (2.5) 220 (1.7)	29 (0.8) 227 (1.5)
45 minutes	1992	3 (1.0) 223 (8.9)!	11 (0.7) 224 (2.5)	3 (1.4) *** (**.*)	12 (1.2) 218 (3.9)	4 (0.9) 202 (4.6)!	12 (0.5) 218 (1.5)
	1996	4 (1.2) 212 (5.6)!	11 (0.9) 223 (2.8)	2 (1.3) *** (**.*)	13 (0.8) 228 (2.4)	4 (0.9) 210 (5.1)!	13 (0.7) 225 (1.5)>
One hour or more	1992	1 (0.4) *** (**.*)	10 (0.7) 214 (2.7)	1 (****) *** (**.*)	11 (1.4) 207 (4.5)	1 (0.4) *** (**.*)	12 (0.7) 205 (1.8)
	1996	2 (0.8) *** (**.*)	11 (0.9) 210 (3.2)	2 (1.5) *** (**.*)	10 (1.2) 210 (6.3)	2 (0.6) 208 (5.0)!	12 (0.6) 209 (1.9)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. --- Does not apply to teachers. ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic. *** Sample size is insufficient to permit a reliable estimate. **** Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

In addition to being asked about mathematics homework, students were asked how often they use a computer at home for schoolwork. The question was not restricted to mathematics homework, so students' reports most likely included homework for other academic areas such as English and science. Given that home computers are steadily assuming more importance in completing homework assignments,[55] it is important that NAEP monitor the prevalence of this practice and its relationship to performance.

Based on the reports of fourth graders in New Jersey, as shown in Table 6.8:

- Less than half of the students reported that there was no computer at home (39 percent) and another 28 percent reported never or hardly ever using their home computer to do homework.
- Less than one fifth of the fourth graders reported using their home computer to do homework almost every day (6 percent) or once or twice a week (10 percent).
- The average scale score for students who used a computer almost every day for homework (225) was not significantly different from* that of students who never or hardly ever did so (232).
- The average scale score for students who used a computer almost every day for homework was higher than that of students who did not have a computer at home (215).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 6.8 — GRADE 4

Public School Students' Reports on Using Computers at Home

How often do you use a computer at home for schoolwork?	New Jersey	Northeast	Nation
	Percentage and Average Scale Score		
There is no computer at home.	39 (1.7) 215 (1.9)	37 (3.2) 216 (3.2)	44 (1.4) 214 (1.3)
Never or hardly ever	28 (1.0) 232 (1.5)	28 (2.7) 234 (2.0)	26 (0.9) 229 (1.3)
Once or twice a month	16 (1.1) 247 (1.9)	14 (1.5) 244 (2.1)	11 (0.6) 239 (1.6)
Once or twice a week	10 (0.8) 230 (2.5)	12 (1.1) 231 (4.9)	10 (0.6) 228 (2.4)
Almost every day	6 (0.6) 225 (3.6)	9 (1.4) 213 (5.9)!	8 (0.5) 214 (2.4)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

* Although the difference may appear large, recall that "significance" here refers to "statistical significance."

[55] National Center for Education Statistics. *Digest of Education Statistics 1995*. (Washington, DC: National Center for Education Statistics, 1995).

Calculator and Computer Use in the Mathematics Classroom

Recommendations for facilitating mathematics instruction in the nation's schools often include more use of calculators and computers.[56] The NCTM *Standards* recognize the technological world in which students are living and the opportunities that technology provides for students to learn and use mathematics. The increasingly technical workplace demands that students have a deep understanding of mathematics that permits solving complex problems.

Given the importance of using technology in mathematics instruction, NAEP asked students and their teachers about their use of calculators and computers. Teachers in New Jersey were also asked about the availability of computers for mathematics instruction.

Calculators

Recent analysis of data from the NAEP 1992 assessment suggests a positive relationship between calculator use and the effectiveness of the school.[57] The same research also found that the use of calculators in mathematics classes varied widely from school to school. As part of the NAEP assessment, students and their mathematics teachers were asked to report on the frequency of use of calculators in mathematics classes and teachers were asked about the use of calculators on tests. This latter question is relevant to NAEP since the students are allowed to use a calculator on a portion of the assessment. (The question concerning calculator use on tests was not asked of students.) Reports from public school teachers of the fourth graders in New Jersey (shown in Tables 6.9 and 6.10) indicate the following:

- Less than half of the students used a calculator in their mathematics class almost every day (6 percent) or once or twice a week (39 percent). Less than one fifth of the students never or hardly ever used a calculator (15 percent). The percentage of students using a calculator almost every day did not differ significantly from that for the nation (5 percent).
- Teachers of 6 percent of fourth graders reported permitting students to use calculators for tests. The average scale score for students who were allowed to use calculators (213) was not significantly different from* that of students who were not (229).
- The percentage of students in 1996 who were permitted to use calculators on tests was not significantly different from that in 1992 (4 percent).

When fourth graders were asked about the use of calculators in their mathematics class, their responses (as shown in Table 6.9) indicated the following:

- Less than half of the students reported using a calculator almost every day (12 percent) or once or twice a week (28 percent). In comparison, 30 percent of students reported never or hardly ever using calculators in their mathematics class.

---

* Although the difference may appear large, recall that "significance" here refers to "statistical significance."

[56] National Council of Teachers of Mathematics. *Curriculum and Evaluation Standards for School Mathematics*. (Reston, VA: NCTM, 1989).

[57] Mullis, I.V.S., F. Jenkins, and E.G. Johnson. *Effective Schools in Mathematics*. (Washington, DC: U.S. Department of Education, 1994).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 6.9 — GRADE 4

Public School Teachers' and Students' Reports on the Frequency of Calculator Use

How often do the students in this class (do you) use a calculator?	New Jersey		Northeast		Nation	
	Teacher	Student	Teacher	Student	Teacher	Student
	Percentage and Average Scale Score					
Never or hardly ever	15 (3.1) 218 (3.3)!	30 (2.2) 223 (1.8)	20 (4.0) 202 (6.6)!	34 (3.0) 226 (2.4)	24 (2.5) 216 (2.3)	40 (1.4) 220 (1.3)
Once or twice a month	40 (3.9) 225 (2.7)	30 (1.7) 236 (1.9)	39 (6.4) 230 (2.9)	28 (1.8) 239 (2.2)	42 (2.6) 223 (1.6)	26 (0.8) 233 (1.0)
Once or twice a week	39 (3.9) 234 (2.6)	28 (1.7) 229 (2.1)	32 (4.7) 238 (3.6)	26 (2.8) 225 (2.8)	29 (2.4) 229 (1.8)	23 (1.0) 223 (1.4)
Almost every day	6 (1.9) 228 (4.9)!	12 (1.1) 215 (3.3)	9 (3.1) *** (**.*)	12 (1.5) 206 (4.4)	5 (0.9) 227 (5.0)	11 (0.6) 206 (1.9)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic. *** Sample size is insufficient to permit a reliable estimate.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 6.10 — GRADE 4

Public School Teachers' Reports on the Use of Calculators for Tests

Do you permit students in this class to use calculators for tests?		New Jersey	Northeast	Nation
		Percentage and Average Scale Score		
Yes	1992	4 (1.5) 242 (9.3)!	9 (4.0) 236 (7.6)!	5 (1.2) 228 (4.3)!
	1996	6 (1.9) 213 (7.8)!	16 (5.8) 225 (4.4)!	10 (1.8) 222 (2.4)
No	1992	96 (1.5) 227 (1.4)	91 (4.0) 220 (2.6)	95 (1.2) 218 (1.0)
	1996	94 (1.9) 229 (1.5)	84 (5.8) 226 (2.8)	90 (1.8) 223 (1.2)>

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

Computers

Computers are potentially valuable instructional tools for the mathematics classroom and can be used to demonstrate mathematics concepts, diagnose learning problems, deliver instruction, and analyze data. Computers are increasingly important in students' homes, where they are used for homework as well as for other pursuits. Since 1984, the percentage of students in grades 7 through 12 who use a computer at school or at home has increased.[58]

Teachers of the students assessed were asked about the availability and accessibility of computers for use in their mathematics classroom. Based on the responses of fourth-grade teachers, the results are shown in Table 6.11.

- In New Jersey, 14 percent of students had teachers who reported that no computers were available for use in their mathematics class and 14 percent had teachers who reported that computers were available in a computer laboratory but difficult to access or schedule. In comparison, 2 percent of students were in mathematics classes where four or more computers were available within the classroom and 15 percent where computers were available in a laboratory and easy to access or schedule.

- The percentage of students in mathematics classes where computers were not available was larger than the percentage for the nation (6 percent).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 6.11 — GRADE 4

Public School Teachers' Reports on the Availability of Computers

Which best describes the availability of computers for use by students in your mathematics classes?	New Jersey	Northeast	Nation
	Percentage and Average Scale Score		
None available	14 (2.7) 214 (4.7)	8 (2.0) *** (**.*)	6 (1.3) 218 (4.0)!
One within the classroom	43 (4.5) 227 (2.4)	28 (4.7) 230 (3.7)!	35 (3.3) 225 (1.7)
Two or three within the classroom	11 (2.7) 240 (5.5)!	25 (2.8) 232 (5.3)	22 (2.2) 222 (2.1)
Four or more within the classroom	2 (1.0) *** (**.*)	7 (1.9) *** (**.*)	7 (1.5) 224 (7.5)!
Available in a computer laboratory but difficult to access or schedule	14 (2.9) 236 (3.4)!	18 (4.2) 218 (9.2)!	13 (1.9) 222 (4.5)
Available in a computer laboratory and easy to access or schedule	15 (3.4) 226 (4.2)!	14 (2.7) 232 (3.4)!	16 (2.7) 224 (2.0)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic. *** Sample size is insufficient to permit a reliable estimate.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

[58] National Center for Education Statistics. *Digest of Education Statistics 1995*. (Washington, DC: National Center for Education Statistics, 1995).

In addition to a range of availability from school to school, the uses of computers can vary widely from class to class. There are a variety of ways that computers can be used to help students learn and use mathematics, such as exploring new mathematical ideas, analyzing information to solve problems, practicing skills, and playing mathematical games. Also, the frequency of use can vary regardless of the primary use of the computers in the classroom. Teachers in New Jersey were asked how they used computers and how often they were used in their mathematics classroom. The responses of fourth-grade public school teachers, shown in Table 6.12, indicate the following:

- About one third of the fourth graders had teachers who reported not using a computer for mathematics instruction (34 percent). This percentage was larger than the percentage for the nation (23 percent).
- In New Jersey, 40 percent of students had teachers who reported never or hardly ever using a computer with their classes, compared to less than half who reported doing so almost every day (5 percent) or once or twice a week (32 percent).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 6.12 — GRADE 4

Public School Teachers' Reports on the Primary Use and Frequency of Use of Computers

	New Jersey	Northeast	Nation
	Percentage and Average Scale Score		
If you do use computers, what is the primary use of these computers for mathematics instruction?			
Drill and practice	30 (4.7) 225 (2.4)	19 (2.9) 229 (3.3)!	27 (2.4) 222 (2.1)
Demonstration of new topics in mathematics	1 (****) *** (**.*)	3 (1.6) *** (**.*)	2 (0.6) 222 (7.5)!
Playing mathematical/learning games	33 (4.2) 234 (2.6)	47 (5.4) 236 (3.5)	41 (2.6) 225 (1.5)
Simulations and applications	2 (1.0) *** (**.*)	7 (3.4) *** (**.*)	6 (1.2) 225 (3.6)
I do not use computers for instruction.	34 (4.2) 225 (3.4)	24 (4.0) 207 (7.6)!	23 (2.9) 220 (3.4)
How often do the students in this class use a computer?			
Never or hardly ever	40 (4.4) 225 (2.6)	18 (3.7) 213 (4.9)!	21 (2.4) 221 (2.6)
Once or twice a month	22 (2.7) 230 (2.2)	17 (4.5) 246 (4.3)!	19 (1.9) 227 (2.1)
Once or twice a week	32 (3.7) 231 (3.0)	46 (4.5) 222 (2.7)	46 (2.5) 222 (1.4)
Almost every day	5 (1.3) 217 (5.8)!	19 (3.1) 231 (3.5)!	14 (1.8) 225 (2.8)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic. *** Sample size is insufficient to permit a reliable estimate. **** Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

Finally, students were asked how often they used computers when doing mathematics in school. The question was not limited to using the computer in their mathematics class. Therefore, students' responses could include use of computers to do mathematics assignments at other times throughout the school day or in before/after school programs. On the basis of the responses of fourth-grade public school students, as shown in Table 6.13, results indicated that:

- In New Jersey, 63 percent of students never or hardly ever used computers to do mathematics in school. About one quarter of the fourth graders used computers for this purpose almost every day (9 percent) or once or twice a week (18 percent).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 6.13 — GRADE 4

Public School Students' Reports on the Frequency of Computer Use

When you do mathematics in school, how often do you use a computer?	New Jersey	Northeast	Nation
	Percentage and Average Scale Score		
Never or hardly ever	63 (1.8) 227 (1.4)	54 (2.0) 228 (2.6)	56 (1.0) 223 (1.2)
Once or twice a month	10 (0.8) 237 (2.8)	10 (2.0) 228 (3.2)!	10 (0.7) 226 (2.4)
Once or twice a week	18 (1.3) 227 (2.2)	20 (2.0) 225 (3.6)	20 (1.0) 225 (1.3)
Almost every day	9 (0.8) 221 (4.3)	15 (1.2) 224 (3.5)	14 (0.8) 218 (1.8)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

CHAPTER 7

Influences Beyond School That Facilitate Learning Mathematics

The home environment can be an important support for the school environment. To examine the relationship between mathematics scale scores and home factors, the NAEP assessment considered students' responses to questions about home factors and principals' responses to questions about parental involvement in the school. To examine the impact of student mobility on academic achievement, students were also asked how often they had changed schools because of household moves.

The students' attitudes toward mathematics can also be expected to relate to their performance in the assessment. As the NCTM *Curriculum and Evaluation Standards for School Mathematics* warns, the beliefs that students develop influence not only their thinking and performance but also their attitude and decisions about studying mathematics in future years. The NCTM *Standards* describes specific attitudes that should be given increased attention in the curriculum.

Discussing Studies at Home

When students discuss academic work at home, they create an important link between home and school. How often schoolwork is discussed at home can be a measure of the importance of school for students and their families. Recent NAEP assessments in a variety of subject areas have found a positive relationship between discussing studies at home and student performance.[59]

[59] Campbell, J.R., P.L. Donahue, C.M. Reese, and G.W. Phillips. *NAEP 1994 Reading Report Card for the Nation and the States*. (Washington, DC: National Center for Education Statistics, 1996); Beatty, A.S., C.M. Reese, H.R. Persky, and P. Carr. *NAEP 1994 U.S. History Report Card*. (Washington, DC: National Center for Education Statistics, 1996); Persky, H.R., C.M. Reese, C.Y. O'Sullivan, S. Lazer, J. Moore, and S. Shakrani. *NAEP 1994 Geography Report Card*. (Washington, DC: National Center for Education Statistics, 1996).

Students were asked to report on the frequency of home discussion about schoolwork. As shown in Table 7.1, the 1996 results for fourth graders attending public schools in New Jersey indicate that

- More than half of the fourth graders (58 percent) said they discussed their schoolwork at home almost every day. This percentage was larger than the percentage who said they never or hardly ever had such discussions (13 percent).

- The average scale score for students who discussed their schoolwork almost every day (228) was higher than that for students who never or hardly ever did so (218).

THE NATION'S REPORT CARD NAEP 1996 State Assessment

TABLE 7.1 — GRADE 4

Public School Students' Reports on Discussing Studies at Home

How often do you discuss things you have studied in school with someone at home?		New Jersey	Northeast	Nation
		Percentage and Average Scale Score		
Never or hardly ever	1992	14 (1.1) 214 (2.5)	19 (1.7) 212 (2.7)	20 (0.6) 209 (1.2)
	1996	13 (0.8) 218 (2.7)	15 (1.5) 212 (2.7)	17 (0.6)< 213 (1.4)
Once or twice a month	1992	4 (0.5) 233 (4.1)	5 (0.7) *** (**.*)	6 (0.3) 220 (2.9)
	1996	5 (0.5) 232 (2.8)	4 (0.5) *** (**.*)	5 (0.3) 226 (2.8)
Once or twice a week	1992	23 (1.1) 232 (1.7)	21 (0.9) 229 (2.9)	21 (0.6) 225 (1.1)
	1996	24 (1.1) 231 (1.8)	22 (2.2) 231 (3.2)	22 (0.7) 227 (1.4)
Almost every day	1992	58 (1.5) 228 (1.7)	55 (1.8) 226 (2.8)	53 (0.7) 220 (0.9)
	1996	58 (1.3) 228 (1.5)	59 (1.9) 228 (3.1)	56 (0.9) 223 (1.3)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. *** Sample size is insufficient to permit a reliable estimate.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

Literacy Materials in the Home

Students can learn much about mathematics and its role in real-world situations by reading materials outside the classroom. Research findings and results from opinion polls are often found in magazine and newspaper articles. Also, the availability of reading and reference materials at home may be an indicator of the value placed by parents on learning.[60] In past NAEP assessments, a positive relationship has consistently been reported between print materials in the home and average scale scores.[61]

As part of the NAEP assessment, students were asked whether their families have more than 25 books, an encyclopedia, receive a newspaper regularly, and receive any magazines regularly. Table 7.2 shows the percentages of fourth-grade public school students reporting that their families have all four types, only three types, or two or fewer types of these literacy materials and students' average scale scores. Based on their responses:

- Less than half of the students in New Jersey (40 percent) reported having all four types of literacy materials in their homes. This percentage was greater than the percentage for the nation (34 percent).

- In comparison, the percentage reporting having two or fewer types of these materials (28 percent) was smaller than the percentage having all four types. The percentage having two or fewer types was smaller than the percentage for the nation (32 percent).

- In 1996 the average mathematics scale score for students with all four types of literacy materials (236) was higher than that for students with two or fewer types (213).

---

[60] Rogoff, B. *Apprenticeship in Thinking: Cognitive Development in Social Context*. (New York: Oxford University Press, 1990).

[61] Campbell, J.R., P.L. Donahue, C.M. Reese, and G.W. Phillips. *NAEP 1994 Reading Report Card for the Nation and the States*. (Washington, DC: National Center for Education Statistics, 1996); Beatty, A.S., C.M. Reese, H.R. Persky, and P. Carr. *NAEP 1994 U.S. History Report Card*. (Washington, DC: National Center for Education Statistics, 1996); Persky, H.R., C.M. Reese, C.Y. O'Sullivan, S. Lazer, J. Moore, and S. Shakrani. *NAEP 1994 Geography Report Card*. (Washington, DC: National Center for Education Statistics, 1996).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 7.2 — GRADE 4

Public School Students' Reports on Literacy Materials in the Home

How many of the following types of reading materials are in your home (more than 25 books, an encyclopedia, a newspaper, magazines)?		New Jersey	Northeast	Nation
		Percentage and Average Scale Score		
Zero to two	1992	24 (1.5) 214 (2.3)	26 (2.8) 207 (2.5)	31 (1.3) 208 (1.1)
	1996	28 (1.7) 213 (1.9)	28 (1.6) 214 (4.0)	32 (0.9) 212 (1.7)
Three	1992	34 (1.1) 227 (1.6)	36 (1.9) 225 (2.3)	35 (0.7) 220 (1.0)
	1996	33 (1.0) 228 (1.8)	35 (1.1) 226 (3.0)	34 (0.7) 224 (1.2)>
Four	1992	42 (1.8) 235 (1.6)	38 (2.5) 234 (3.1)	34 (1.2) 228 (1.2)
	1996	40 (1.7) 236 (1.4)	37 (1.9) 236 (1.6)	34 (0.9) 231 (1.3)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

Television Viewing Habits

Past NAEP assessments have shown that over 40 percent of fourth- and eighth-grade students reported watching four or more hours of television each day. A major concern is that time spent watching television results in less time available for homework and related academic activities. The effects of such extensive television exposure are difficult to document, but there is a generally negative relationship between NAEP score results and hours watched.[62]

Students were asked how much television they usually watched each day. The results for fourth-grade public school students in New Jersey are shown in Table 7.3.

- Among fourth graders, 22 percent reported watching six or more hours of television on a typical day. This percentage did not differ significantly from the percentage who reported watching one hour or less (19 percent).

- The percentage of fourth graders in New Jersey who reported watching six or more hours of television a day was not significantly different from the percentage for the nation (20 percent).

- Based on the 1996 state results, the average mathematics scale score for fourth-grade students who reported watching two to three hours of television a day (234) was not significantly different from that for students who reported watching one hour or less (231).

- The average scale score for fourth graders who reported watching two to three hours of television a day was higher than that for students who reported watching six hours or more (212).

[62] Campbell, J.R., P.L. Donahue, C.M. Reese, and G.W. Phillips. *NAEP 1994 Reading Report Card for the Nation and the States*. (Washington, DC: National Center for Education Statistics, 1996); Beatty, A.S., C.M. Reese, H.R. Persky, and P. Carr. *NAEP 1994 U.S. History Report Card*. (Washington, DC: National Center for Education Statistics, 1996); Persky, H.R., C.M. Reese, C.Y. O'Sullivan, S. Lazer, J. Moore, and S. Shakrani. *NAEP 1994 Geography Report Card*. (Washington, DC: National Center for Education Statistics, 1996).

TABLE 7.3 — GRADE 4

Public School Students' Reports on Television Viewing Habits

How much television do you usually watch each day?		New Jersey	Northeast	Nation
		Percentage and Average Scale Score		
1 hour or less	1992	20 (1.2) 231 (2.5)	20 (1.9) 230 (3.2)	21 (0.8) 221 (1.6)
	1996	19 (1.1) 231 (2.2)	23 (1.9) 224 (3.9)	24 (1.1) 223 (1.8)
2 to 3 hours	1992	38 (1.3) 233 (1.7)	36 (1.8) 231 (3.0)	35 (0.8) 225 (1.1)
	1996	35 (1.3) 234 (1.8)	37 (2.2) 233 (2.0)	36 (0.7) 228 (1.2)
4 to 5 hours	1992	22 (1.2) 228 (1.6)	22 (2.5) 224 (2.5)	22 (0.8) 221 (1.3)
	1996	24 (1.3) 229 (1.7)	22 (1.3) 232 (3.0)	20 (0.7) 224 (1.5)
6 hours or more	1992	20 (1.3) 212 (2.2)	23 (2.6) 206 (3.2)	22 (0.8) 204 (1.1)
	1996	22 (1.4) 212 (2.3)	18 (2.0) 208 (3.1)	20 (0.8) 208 (1.5)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

Parental Support

When parents are involved in their children's education, both children and parents are likely to benefit. Research on students at risk has shown that parents' participation in their children's education has more effect on the child's performance than parent income or parent education.[63] Parental involvement is naturally part of the home environment, but it is also increasingly sought in the school.

As part of the NAEP assessment, the principals of participating students were asked about parental involvement in their schools. Table 7.4 presents the results for fourth graders in public schools in New Jersey. According to these results:

- Overall, almost all of the fourth-grade students attended schools where principals characterized parental support as very positive (55 percent) or somewhat positive (41 percent).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 7.4 — GRADE 4

Public Schools' Reports on Parental Support

How would you characterize parental support for student achievement within your school?		New Jersey	Northeast	Nation
		Percentage and Average Scale Score		
Somewhat to very negative	1992	5 (2.1) 204 (10.1)!	1 (****) *** (**.*)	5 (1.2) 202 (4.9)!
	1996	4 (2.0) *** (**.*)	1 (****) *** (**.*)	5 (1.5) 200 (5.8)!
Somewhat positive	1992	39 (4.5) 218 (2.8)	47 (6.9) 216 (3.0)	54 (3.2) 215 (1.3)
	1996	41 (5.3) 218 (2.5)	55 (7.6) 220 (3.0)	52 (3.7) 220 (1.3)>
Very positive	1992	56 (4.4) 236 (1.9)	52 (7.5) 231 (4.0)	41 (3.2) 225 (1.7)
	1996	55 (5.2) 237 (2.3)	44 (8.2) 238 (2.7)	43 (3.9) 227 (1.7)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic. *** Sample size is insufficient to permit a reliable estimate. **** Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

[63] Office of Educational Research and Improvement. *Mapping out the National Assessment of Title I: The Interim Report — 1996*. (Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education, 1996).

Student Mobility

Research indicates that moving more than once or twice during the school year lowers student achievement. Students who attend the same school throughout their career are most likely to graduate, whereas the most mobile of the school populations have the highest rates of failure and dropping out.[64]

In order to look at the relationship between mobility and mathematics achievement, students were asked how many times within the past two years they had changed schools because they had changed where they lived. Table 7.5 shows results for fourth-grade public school students.

- In terms of student mobility, 66 percent of fourth graders reported not moving over the last two years while 8 percent of students reported doing so three or more times. The students with the highest reported mobility had an average scale score (204) that was lower than that of those who reported not moving (233).
- The percentage of students in New Jersey who reported moving three or more times (8 percent) was smaller than the percentage for the nation (11 percent).

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE 7.5 — GRADE 4

Public School Students' Reports on Mobility

Within the past two years, how many times have you changed schools because you changed where you lived?		New Jersey	Northeast	Nation
		Percentage and Average Scale Score		
None	1992	69 (1.3) 233 (1.4)	65 (1.9) 230 (2.2)	60 (0.9) 225 (0.9)
	1996	66 (1.4) 233 (1.4)	66 (2.6) 231 (2.5)	62 (1.1) 227 (1.2)
One	1992	17 (0.9) 221 (1.8)	19 (1.7) 219 (4.6)	19 (0.6) 216 (1.6)
	1996	18 (1.2) 223 (2.4)	18 (1.6) 225 (5.0)	19 (0.7) 221 (1.9)
Two	1992	6 (0.6) 208 (3.9)	7 (0.9) 209 (5.2)	9 (0.4) 206 (1.7)
	1996	8 (0.6) 209 (3.4)	9 (1.4) 212 (3.5)!	8 (0.5) 210 (2.1)
Three or more	1992	8 (0.8) 207 (2.7)	10 (1.3) 204 (4.7)	13 (0.5) 201 (1.6)
	1996	8 (0.7) 204 (3.3)	7 (0.7) 209 (5.1)	11 (0.8) 207 (2.0)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. ! Interpret with caution — the nature of the sample does not allow accurate determination of the variability of this statistic.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

---

[64] ERIC Clearinghouse on Urban Education. *Highly Mobile Students: Educational Problems and Possible Solutions*. (New York, NY: ERIC Clearinghouse on Urban Education, ERIC/CUE Digest, Number 73, 1991).

Students' Views About Mathematics

The attitudes children form about mathematics can affect the depth to which they learn the concepts. These same attitudes can also affect decisions that middle school students make about what mathematics courses they will study. Failure to study mathematics can close the doors to education beyond high school, and to many interesting and exciting careers. Thus, students' attitudes and beliefs about mathematics can be a contributing factor affecting the skills they will acquire.

Do Students Believe That Math Is Useful for Everyday Problems?

If children view mathematics as a practical, useful subject, they may better understand that it can be applied to a wide variety of real-world problems and phenomena. The NCTM *Standards* explain that, even though most mathematical ideas in the kindergarten through fourth grade curriculum arise *from* the everyday world, they must be regularly applied *to* real-world situations. Further, children need to understand that mathematics is an integral part of real-world situations and activities in other curricular areas. One major purpose of mathematics is to help children understand and interpret their world and solve problems that occur in it. This important view of mathematics must continue through the curriculum for grades 5 through 8. Teachers should emphasize the application of mathematics to real-world problems as well as to other settings relevant to middle school students.

In order to examine whether mathematics has been made relevant to the students, they were asked the degree to which they agreed or disagreed with the statement that mathematics is useful for solving everyday problems. Responses by fourth-grade public school students are reported in Table 7.6. According to their responses:

- About two thirds of the fourth graders in New Jersey agreed with the statement that mathematics is useful for solving everyday problems (69 percent). This percentage was greater than that of students disagreeing with the statement (13 percent).

- The percentage of students in the state agreeing that mathematics is useful for everyday problems (69 percent) was not significantly different from the percentage seen nationally (68 percent).

TABLE 7.6 — GRADE 4

Public School Students' Views on the Usefulness of Mathematics

To what degree do you agree with the statement "Mathematics is useful for solving everyday problems?"		New Jersey	Northeast	Nation
		Percentage and Average Scale Score		
Disagree	1992	13 (0.8) 220 (2.6)	12 (1.1) 209 (3.6)	13 (0.6) 207 (1.7)
	1996	13 (0.8) 215 (2.3)	12 (0.6) 216 (4.1)	14 (0.7) 211 (2.0)
Undecided	1992	22 (1.0) 225 (2.0)	21 (1.7) 220 (2.5)	21 (0.9) 218 (1.3)
	1996	18 (0.8)< 223 (2.2)	16 (0.9) 223 (2.0)	18 (0.7)< 221 (1.5)
Agree	1992	65 (1.3) 231 (1.5)	67 (2.0) 230 (2.3)	66 (1.1) 223 (0.9)
	1996	69 (1.1)> 232 (1.5)	71 (1.1) 232 (2.8)	68 (0.9) 228 (1.0)>

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

Do Students Believe Mathematics Is a Static Discipline?

Do students believe that mathematics is a static, unchanging, rule-bound discipline or a dynamic, flexible way of approaching problem-solving situations? This question is key to the mathematics curricula described in the NCTM *Standards*. Curricula should emphasize the development of students' mathematical thinking and reasoning abilities. Although learning the basic facts and rules remains important, memorization of facts and rules without understanding and not being able to use them appropriately is not helpful. Curricula should also emphasize the importance of flexibility in choosing strategies and techniques for solving mathematical problems. Successful problem solving, employing flexibility in approach and technique, should lead to confidence and perseverance in solving higher level problems.

Students were asked whether they agreed or disagreed with the two statements that learning mathematics is mostly memorizing facts and that there is only one way to solve a mathematical problem. Responses by fourth-grade public school students are reported in Table 7.7. According to their responses, the following is true:

- About one quarter of the fourth graders in New Jersey disagreed with the statement that mathematics is mostly memorizing facts (23 percent). This percentage was smaller than that of students agreeing with the statement (53 percent).

- When asked if there is only one way to solve a mathematics problem, 69 percent of fourth graders disagreed. This percentage was greater than that of students agreeing with this belief (13 percent).

- The percentage of students in the state disagreeing that mathematics is the memorization of facts (23 percent) was not significantly different from the percentage in the nation (21 percent). However, the percentage disagreeing with the belief that there is only one solution to a mathematics problem (69 percent) was greater than the national percentage (63 percent).

TABLE 7.7 — GRADE 4

Public School Students' Views on the Nature of Mathematics

To what degree do you agree with the following statements?	New Jersey	Northeast	Nation
	Percentage and Average Scale Score		
Learning mathematics is mostly memorizing facts.			
Disagree	23 (1.0) 237 (2.4)	25 (1.9) 239 (2.7)	21 (0.9) 233 (1.6)
Undecided	24 (1.0) 234 (1.8)	26 (1.2) 231 (3.2)	25 (0.6) 226 (1.3)
Agree	53 (1.5) 222 (1.4)	49 (2.5) 221 (2.8)	55 (0.9) 220 (1.0)
There is only one correct way to solve a mathematics problem.			
Disagree	69 (1.4) 234 (1.3)	70 (1.4) 234 (1.9)	63 (0.9) 231 (1.0)
Undecided	17 (1.1) 222 (2.2)	18 (1.4) 221 (4.6)	20 (0.7) 219 (1.7)
Agree	13 (0.9) 206 (2.9)	13 (1.1) 204 (5.3)	17 (0.6) 205 (1.6)

The NAEP mathematics scale ranges from 0 to 500. The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details).
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

APPENDIX A

Reporting NAEP 1996 Mathematics Results for New Jersey

A.1 Participation Guidelines

As was discussed in the Introduction, unless the overall participation rate is sufficiently high for a jurisdiction, there is a risk that the assessment results for that jurisdiction will be subject to appreciable nonresponse bias. Moreover, even if the overall participation rate is high, there may be significant nonresponse bias if the nonparticipation that does occur is heavily concentrated among certain types of schools or students. The following guidelines concerning school and student participation rates in the state assessment program were established to address four significant ways in which nonresponse bias could be introduced into the jurisdiction sample estimates. The guidelines determining a jurisdiction's eligibility to have its results published are presented below. Also presented below are the conditions that will result in a jurisdiction's receiving a notation in the 1996 reports. Note that in order for a jurisdiction's results to be published with no notations, that jurisdiction must satisfy all guidelines. (A more complete discussion of the NAEP participation guidelines can be found in the *Technical Report of the NAEP 1996 State Assessment Program in Mathematics*.)

Guidelines on the Publication of NAEP Results

Guideline 1 — Publication of Public School Results

A jurisdiction will have its public school results published in the *NAEP 1996 Mathematics Report Card* (or in other reports that include all state-level results) if and only if its weighted participation rate for the initial sample of public schools is greater than or equal to 70 percent. Similarly, a jurisdiction will receive a separate *NAEP 1996 Mathematics State Report* if and only if its weighted participation rate for the initial sample of public schools is greater than or equal to 70 percent.

Guideline 2 — Publication of Nonpublic School Results

A jurisdiction will have its nonpublic school results published in the *NAEP 1996 Mathematics Report Card* (or in other reports that include all state-level results) if and only if its weighted participation rate for the initial sample of nonpublic schools is greater than or equal to 70 percent **AND** meets minimum sample size requirements.[1] A jurisdiction eligible to receive a separate *NAEP 1996 Mathematics State Report* under guideline 1 will have its nonpublic school results included in that report if and only if that jurisdiction's weighted participation rate for the initial sample of nonpublic schools is greater than or equal to 70 percent **AND** meets minimum sample size requirements. If a jurisdiction meets guideline 2 but fails to meet guideline 1, a separate *NAEP 1996 Mathematics State Report* will be produced containing only nonpublic school results.

Guideline 3 — Publication of Combined Public and Nonpublic School Results

A jurisdiction will have its combined results published in the *NAEP 1996 Mathematics Report Card* (or in other reports that include all state-level results) if and only if both guidelines 1 and 2 are satisfied. Similarly, a jurisdiction eligible to receive a separate *NAEP 1996 Mathematics State Report* under guideline 1 will have its combined results included in that report if and only if guideline 2 is also met.

Guidelines for Notations of NAEP Results

Guideline 4 — Notation for Overall Public School Participation Rate

A jurisdiction that meets guideline 1 will receive a notation if its weighted participation rate for the initial sample of public schools was below 85 percent **AND** the weighted public school participation rate after substitution was below 90 percent.

Guideline 5 — Notation for Overall Nonpublic School Participation Rate

A jurisdiction that meets guideline 2 will receive a notation if its weighted participation rate for the initial sample of nonpublic schools was below 85 percent **AND** the weighted nonpublic school participation rate after substitution was below 90 percent.

[1] Minimum participation size requirements for reporting nonpublic school data consist of two components: (1) a school sample size of six or more participating schools and (2) an assessed student sample size of at least 62.

Guideline 6 — Notation for Strata-Specific Public School Participation Rate

A jurisdiction that is not already receiving a notation under guideline 4 will receive a notation if the sample of public schools included a class of schools with similar characteristics that had a weighted participation rate (after substitution) of below 80 percent, and from which the nonparticipating schools together accounted for more than five percent of the jurisdiction's total weighted sample of public schools. The classes of schools from each of which a jurisdiction needed minimum school participation levels were determined by degree of urbanization, minority enrollment, and median household income of the area in which the school is located.

Guideline 7 — Notation for Strata-Specific Nonpublic School Participation Rate

A jurisdiction that is not already receiving a notation under guideline 5 will receive a notation if the sample of nonpublic schools included a class of schools with similar characteristics that had a weighted participation rate (after substitution) of below 80 percent, and from which the nonparticipating schools together accounted for more than five percent of the jurisdiction's total weighted sample of nonpublic schools. The classes of schools from each of which a jurisdiction needed minimum school participation levels were determined by type of nonpublic school (Catholic versus non-Catholic) and location (metropolitan versus nonmetropolitan).

Guideline 8 — Notation for Overall Student Participation Rate in Public Schools

A jurisdiction that meets guideline 1 will receive a notation if the weighted student response rate within participating public schools was below 85 percent.

Guideline 9 — Notation for Overall Student Participation Rate in Nonpublic Schools

A jurisdiction that meets guideline 2 will receive a notation if the weighted student response rate within participating nonpublic schools was below 85 percent.

Guideline 10 — Notation for Strata-Specific Student Participation Rates in Public Schools

A jurisdiction that is not already receiving a notation under guideline 8 will receive a notation if the sampled students within participating public schools included a class of students with similar characteristics that had a weighted student response rate of below 80 percent, and from which the nonresponding students together accounted for more than five percent of the jurisdiction's weighted assessable public school student sample. Student groups from which a jurisdiction needed minimum levels of participation were determined by the age of the student, whether or not the student was classified as a student with a disability (SD) or of limited English proficiency (LEP), and the type of assessment session (monitored or unmonitored), as well as school level of urbanization, minority enrollment, and median household income of the area in which the school is located.

Guideline 11 — Notation for Strata-Specific Student Participation Rates in Nonpublic Schools

A jurisdiction that is not already receiving a notation under guideline 9 will receive a notation if the sampled students within participating nonpublic schools included a class of students with similar characteristics that had a weighted student response rate of below 80 percent, and from which the nonresponding students together accounted for more than five percent of the jurisdiction's weighted assessable nonpublic school student sample. Student groups from which a jurisdiction needed minimum levels of participation were determined by the age of the student, whether or not the student was classified as a student with a disability (SD) or of limited English proficiency (LEP), and the type of assessment session (monitored or unmonitored), as well as type and location of school.

A.2 NAEP Reporting Groups

The state assessment program provides results for groups of students defined by shared characteristics — region of the country, gender, race/ethnicity, parental education, location of the school, type of school, participation in Title I programs, and eligibility for the free/reduced-price lunch component of the National School Lunch Program. Based on criteria described later in this appendix, results are reported for subpopulations only when sufficient numbers of students and adequate school representation are present. For public school students, the minimum requirement is at least 62 students in a particular subgroup from at least 5 primary sampling units (PSUs).[2] For nonpublic school students, the minimum requirement is 62 students from at least 6 different schools for the state assessment program or from at least 5 PSUs for the national assessment. However, the data for all students, regardless of whether their subgroup was reported separately, were included in computing overall results. Definitions of the subpopulations referred to in this report are presented on the following pages.

Region

Results are reported for four regions of the nation: Northeast, Southeast, Central, and West. States included in each region are shown in Figure A.1. All 50 states and the District of Columbia are listed. Territories and the two Department of Defense Educational Activities jurisdictions were not assigned to any region.

Regional results are based on national assessment samples, not on aggregated state assessment program samples. Thus, the regional results are based on a sample that is different and separate from that used to report the state results.

THE NATION'S REPORT CARD naep 1996 State Assessment

FIGURE A.1

Regions of the Country

NORTHEAST	SOUTHEAST	CENTRAL	WEST
Connecticut	Alabama	Illinois	Alaska
Delaware	Arkansas	Indiana	Arizona
District of Columbia	Florida	Iowa	California
Maine	Georgia	Kansas	Colorado
Maryland	Kentucky	Michigan	Hawaii
Massachusetts	Louisiana	Minnesota	Idaho
New Hampshire	Mississippi	Missouri	Montana
New Jersey	North Carolina	Nebraska	Nevada
New York	South Carolina	North Dakota	New Mexico
Pennsylvania	Tennessee	Ohio	Oklahoma
Rhode Island	Virginia*	South Dakota	Oregon
Vermont	West Virginia	Wisconsin	Texas
Virginia*			Utah
			Washington
			Wyoming

* The part of Virginia that is included in the Washington, DC, metropolitan area is included in the Northeast region; the remainder of the state is in the Southeast region.

[2] For the State Assessment Program, a PSU is most often a single school; for the national assessment, a PSU is a selected geographic region (a county, group of counties, or a metropolitan statistical area).

Gender

Results are reported separately for males and females.

Race/Ethnicity

The race/ethnicity variable is derived from two questions asked of students and schools' records, and it is used for race/ethnicity subgroup comparisons. Two questions from the set of general student background questions were used to determine race/ethnicity:

If you are Hispanic, what is your Hispanic background?

- ° I am not Hispanic.
- ° Mexican, Mexican American, or Chicano
- ° Puerto Rican
- ° Cuban
- ° Other Spanish or Hispanic background

Students who responded to this question by filling in the second, third, fourth, or fifth oval were considered Hispanic. For students who filled in the first oval, did not respond to the question, or provided information that was illegible or could not be classified, responses to the question below were examined in an effort to determine race/ethnicity.

Which best describes you?

- ° White (not Hispanic)
- ° Black (not Hispanic)
- ° Hispanic ("Hispanic" means someone who is from a Mexican, Mexican American, Chicano, Puerto Rican, Cuban, or other Spanish or Hispanic background.)
- ° Asian or Pacific Islander ("Asian or Pacific Islander" means someone who is from a Chinese, Japanese, Korean, Filipino, Vietnamese, or other Asian or Pacific Island background.)
- ° American Indian or Alaskan Native ("American Indian or Alaskan Native" means someone who is from one of the American Indian tribes, or one of the original people of Alaska.)
- ° Other (specify) ______________________

Students' race/ethnicity was then assigned on the basis of their response. For students who filled in the sixth oval ("Other") or provided illegible information or information that could not be classified, or did not respond at all, race/ethnicity was assigned as determined by school records.[3]

Race/ethnicity could not be determined for students who did not respond to either of the demographic questions and whose schools did not provide information about race/ethnicity.

The details of how race/ethnicity classifications were derived is presented so that readers can determine how useful the results are for their particular purposes. Also, some students indicated that they were from a Hispanic background (e.g., Puerto Rican or Cuban) and that a racial/ethnic category other than Hispanic best described them. These students were classified as Hispanic based on the rules described above. Furthermore, information from the schools did not always correspond to how students described themselves. Therefore, the racial/ethnic results presented in this report attempt to provide a clear picture based on several sources of information.

Parents' Highest Level of Education

The variable representing level of parental education is derived from responses to two questions from the set of general student background questions. Students were asked to indicate the extent of their mother's education:

How far in school did your mother go?

- She did not finish high school.
- She graduated from high school.
- She had some education after high school.
- She graduated from college.
- I don't know.

Students were asked a similar question about their father's education level:

How far in school did your father go?

- He did not finish high school.
- He graduated from high school.
- He had some education after high school.
- He graduated from college.
- I don't know.

[3] The procedure for assigning race/ethnicity was modified for Hawaii. See the *Technical Report for the NAEP 1996 State Assessment Program in Mathematics* for details.

The information was combined into one parental education reporting variable determined through the following process. If a student indicated the extent of education for only one parent, that level was included in the data. If a student indicated the extent of education for both parents, the higher of the two levels was included in the data. If a student did not know the level of education for both parents or did not know the level for one parent and did not respond for the other, the parental education level was classified as "I don't know." If the student did not respond for either parent, the student was recorded as having provided no response. (Nationally, 36 percent of fourth graders and 11 percent of eighth graders reported that they did not know the education level of either of their parents.)

Type of Location

Results are provided for students attending public schools in three mutually exclusive location types — central city, urban fringe/large town, and rural/small town — as defined below. The type of location variable is defined in such a way as to indicate the *geographical location* of a student's school. The intention is not to indicate, or imply, social or economic meanings for these location types. The type of location variable, on which the current NAEP sampling is based, does not support the reporting of regional results. Therefore, only state and national results will be presented.

> *Central City*: The Central City category includes central cities of all Metropolitan Statistical Areas (MSAs).[4] Central City is a geographic term and is not synonymous with "inner city."
>
> *Urban Fringe/Large Town*: An Urban Fringe includes all densely settled places and areas within MSAs that are classified as urban by the Bureau of the Census. A Large Town is defined as places outside MSAs with a population greater than or equal to 25,000.
>
> *Rural/Small Town*: Rural includes all places and areas with a population of less than 2,500 that are classified as rural by the Bureau of the Census. A Small Town is defined as places outside MSAs with a population of less than 25,000 but greater than or equal to 2,500.

[4] Each Metropolitan Statistical Area (MSA) is defined by the Office of Management and Budget.

Type of School

Samples for the 1996 state assessment program were expanded to include students attending nonpublic schools (Catholic schools and other religious and private schools) in addition to students attending public schools. The expanded coverage was instituted for the first time in 1994. Samples for the 1990 and 1992 Trial State Assessment programs had been restricted to public school students only. For those jurisdictions meeting pre-established participation rate standards (see earlier section of this appendix), separate results are reported for public schools, for nonpublic schools, and for the combined public and nonpublic school samples. The combined sample for each jurisdiction also contains students attending Bureau of Indian Affairs (BIA) schools and Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) in that jurisdiction. These two categories of schools are not included in either the public or nonpublic school samples.

Note that the DDESS and Department of Defense Dependents Schools (DoDDS)[5] were assessed in 1996 as separate jurisdictions, reported as jurisdictions with public school samples only.

Title I Participation

Based on available school records, students were classified as either currently participating in a Title I program or receiving Title I services, or as not receiving such services. The classification applies only to the school year when the assessment was administered (i.e., the 1995-96 school year) and is not based on participation in previous years. If the school did not offer any Title I programs or services, all students in that school were classified as not participating.

Eligibility for the Free/Reduced-Price School Lunch Program

Based on available school records, students were classified as either currently eligible for the free/reduced-price lunch component of the Department of Agriculture's National School Lunch Program or not eligible. The classification refers only to the school year when the assessment was administered (i.e., the 1995-96 school year) and is not based on eligibility in previous years. If school records were not available, the student was classified as "Information not available." If the school did not participate in the program, all students in that school were classified as "Information not available."

A.3 Guidelines for Analysis and Reporting

This report describes mathematics performance for fourth graders and compares the results for various groups of students within these populations — for example, those who have certain demographic characteristics or who responded to a specific background question in a particular way. The report examines the results for individual demographic groups and individual background questions. It does not include an analysis of the relationships among combinations of these subpopulations or background questions.

[5] The Department of Defense Dependents Schools (DoDDS) refers to overseas schools (i.e., schools outside the United States). Department of Defense Domestic Dependent Elementary and Secondary Schools (DDESS) refers to domestic schools (i.e., schools in the United States).

Drawing Inferences from the Results

Because the percentages of students in these subpopulations and their average scale scores are based on samples — rather than on the entire population of fourth graders in a jurisdiction — the numbers reported are necessarily *estimates*. As such, they are subject to a measure of uncertainty, reflected in the *standard error* of the estimate. When the percentages or average scale scores of certain groups are compared, it is essential to take the standard error into account, rather than to rely solely on observed similarities or differences. Therefore, the comparisons discussed in this report are based on *statistical tests* that consider both the magnitude of the difference between the averages or percentages and the standard errors of those statistics.

One of the goals of the state assessment program is to estimate scale score distributions and percentages of students in the categories described in A.2 for the overall populations of fourth- and eighth-grade students in each participating jurisdiction based on the particular samples of students assessed. The use of *confidence intervals*, based on the standard errors, provides a way to make inferences about the population average scale scores and percentages in a manner that reflects the uncertainty associated with the sample estimates. An estimated sample average scale score ± 2 standard errors approximates a *95 percent confidence interval* for the corresponding population average or percentage. This means that one can conclude with approximately 95 percent confidence that the average scale score of the entire population of interest (e.g., all fourth-grade students in public schools in a jurisdiction) is within ± 2 standard errors of the sample average.

As an example, suppose that the average mathematics scale score of the students in a particular jurisdiction's eighth-grade sample were 256 with a standard error of 1.2. A 95 percent confidence interval for the population average would be as follows:

$$\text{Mean} \pm 2 \text{ standard errors} = 256 \pm 2 \times (1.2) = 256 \pm 2.4 =$$

$$256 - 2.4 \text{ and } 256 + 2.4 = (253.6, 258.4)$$

Thus, one can conclude with 95 percent confidence that the average scale score for the entire population of eighth-grade students in public schools in that jurisdiction is between 253.6 and 258.4.

Similar confidence intervals can be constructed for percentages, *if the percentages are not extremely large or extremely small*. For extreme percentages, confidence intervals constructed in the above manner may not be appropriate, and accurate confidence intervals can be constructed only by using procedures that are quite complicated.

Extreme percentages, defined by both the magnitude of the percentage and the size of the sample from which it was derived, should be interpreted with caution. (The forthcoming *Technical Report of the NAEP 1996 State Assessment Program in Mathematics* contains a more complete discussion of extreme percentages.)

Analyzing Subgroup Differences in Averages and Percentages

The statistical tests determine whether the evidence — based on the data from the groups in the *sample* — is strong enough to conclude that the averages or percentages are really different for those groups in the *population*. If the evidence is strong (i.e., the difference is statistically significant), the report describes the group averages or percentages as being different (e.g., one group performed *higher than* or *lower than* another group) — regardless of whether the sample averages or sample percentages appear to be about the same or not. If the evidence is not sufficiently strong (i.e., the difference is not statistically significant), the averages or percentages are described as being *not significantly different* — again, regardless of whether the sample averages or sample percentages appear to be about the same or widely discrepant. The reader is cautioned to rely on the results of the statistical tests — rather than on the apparent magnitude of the difference between sample averages or percentages — to determine whether those sample differences are likely to represent actual differences between the groups in the population.

In addition to the overall results, this report presents outcomes separately for a variety of important subgroups. Many of these subgroups are defined by shared characteristics of students, such as their gender or race/ethnicity and the type of location in which their school is situated. Other subgroups are defined by the responses of the assessed students' mathematics teachers to questions in the mathematics teacher questionnaire.

In Chapter 1 of this report, differences between the jurisdiction and the nation were tested for overall mathematics scale score and for each of the mathematics content areas. In Chapter 2, significance tests were conducted for the overall scale score for each of the subpopulations. Chapter 3 reports differences between the jurisdiction and nation for the percentage of students at or above the *Proficient* level, and Chapter 4 contains significance tests for the percentage of students at or above the *Proficient* level for each of the subpopulations. In Chapters 5 through 7, comparisons were made across subgroups for responses to various background questions.

As an example of comparisons across subgroups, consider the question: *Do students who reported discussing studies at home almost every day exhibit higher average mathematics scale scores than students who report never or hardly ever doing so?*

To answer the question posed above, begin by comparing the average mathematics scale score for the two groups being analyzed. If the average for the group that reported discussing their studies at home almost every day is higher, it may be tempting to conclude that that group does have a higher mathematics scale score than the group that reported never or hardly ever discussing their studies at home. However, even though the averages differ, there may be no real difference in performance between the two groups in the population because of the uncertainty associated with the estimated average scale scores of the groups in the sample. Remember that the intent is to make a statement about the entire population, not about the particular sample that was assessed. The data from the sample are used to make inferences about the population as a whole.

As discussed in the previous section, each estimated sample average scale score (or percentage) has a degree of uncertainty associated with it. It is therefore possible that if all students in the population (rather than a sample of students) had been assessed or if the assessment had been repeated with a different sample of students or a different, but equivalent, set of questions, the performances of various groups would have been different. Thus, to determine whether there is a *real* difference between the average scale score (or percentage of a certain attribute) for two groups in the population, an estimate of the degree of uncertainty associated with the difference between the scale score averages or percentages of those groups must be obtained 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.

In a manner similar to that in which the standard error for an individual group average or percentage is used, the *standard error of the difference* can be used to help determine whether differences between groups in the population are real. The difference between the mean scale score or percentage 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.

As another example, to determine whether the average mathematics scale score of fourth-grade males is higher than that of fourth-grade females in a particular jurisdiction's public schools, suppose that the sample estimates of the average scale scores and standard errors for males and females were as follows:

Group	Average Scale Score	Standard Error
Males	218	0.9
Females	216	1.1

The difference between the estimates of the average sale scores of males and females is two points (218 - 216). The standard error of this difference is

$$\sqrt{0.9^2 + 1.1^2} = 1.4$$

Thus, an approximate 95 percent confidence interval for this difference is

Mean difference ± 2 standard errors of the difference =

$$2 \pm 2 \times (1.4) = 2 \pm 2.8 = 2 - 2.8 \text{ and } 2 + 2.8 = (-0.8, 4.8)$$

The value zero is within this confidence interval, which extends from -0.8 to 4.8 (i.e., zero is between -0.8 and 4.8). Thus, there is insufficient evidence to claim a difference in average mathematics scale score between the populations of fourth-grade males and females in public schools in the hypothetical jurisdiction.[6]

Throughout this report, when the average scale scores or percentages for two groups were compared, procedures like the one described above were used to draw the conclusions that are presented. If a statement appears in the report indicating that a particular group had a *higher* (or *lower*) average scale score than a second group, the 95 percent confidence interval for the difference between groups did not contain zero. An attempt was made to distinguish between group differences that were statistically significant but rather small in a practical sense and differences that were both statistically and practically significant. A procedure based on effect sizes was used. Statistically significant differences that are rather small are described in the text as *somewhat higher* or *somewhat lower*. When a statement indicates that the average scale score or percentage of some attribute was *not significantly different* for two groups, the confidence interval included zero, and thus no difference could be assumed between the groups. The information described in this section also pertains to comparisons across years. The reader is cautioned to avoid drawing conclusions solely on the basis of the magnitude of the difference. A difference between two groups in the sample that appears to be slight may represent a statistically significant difference in the population because of the magnitude of the standard errors. Conversely, a difference that appears to be large may not be statistically significant.

The procedures described in this section, and the certainty ascribed to intervals (e.g., a 95% confidence interval), are based on statistical theory that assumes that only one confidence interval or test of statistical significance is being performed. However, in each chapter of this report, many different groups are being compared (i.e., multiple sets of confidence intervals are being calculated). In sets of confidence intervals, 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. To hold the certainty level for the set of comparisons at a particular level (e.g., 0.95), adjustments (called multiple comparison procedures) must be made to the methods described in the previous section. One such procedure — the *Bonferroni* method — was used in the analyses described in this report to form confidence intervals for the differences between groups whenever sets of comparisons were considered.[7] Thus, the confidence intervals in the text that are based on sets of comparisons are more conservative than those described on the previous pages.

---

[6] The procedure described above (especially the estimation of the standard error of the difference) is, in a strict sense, only appropriate when the statistics being compared come from independent samples. For certain comparisons in the report, the groups were not independent. In those cases, a different (and more appropriate) estimate of the standard error of the difference was used.

[7] Miller, R.G. *Simultaneous Statistical Inference*. (New York, NY: Wiley, 1966).

Most of the multiple comparisons in this report pertain to relatively small sets or "families" of comparisons. For example, when comparisons were discussed concerning students' reports of parental education, six comparisons were conducted — all pairs of the four parental education levels. In these situations, Bonferroni procedures were appropriate. However, the maps in Chapter 1 of this report display comparisons between New Jersey and all other participating jurisdictions. The "family" of comparisons in this case was as many as 46. To control the certainty level for a large family of comparisons, the False Discovery rate (FDR) criterion[8] was used. Unlike the Bonferroni procedures which control the familywise error rate (i.e., the probability of making even one false rejection in the set of comparisons), the Benjamini and Hochberg (BH) approach using the FDR criterion controls the expected proportion of falsely rejected hypotheses as a proportion of all rejected hypotheses. Bonferroni procedures may be considered conservative for large families of comparisons.[9] In other words, using the Bonferroni method would produce more statistically nonsignificant comparisons than using the BH approach. Therefore, the BH approach is potentially more powerful for comparing New Jersey to all other participating jurisdictions. A more detailed description of the Bonferroni and BH procedures appears in the *Technical Report of the NAEP 1996 State Assessment Program in Mathematics*.

Statistics with Poorly Estimated Standard Errors

Not only are the averages and percentages reported in NAEP subject to uncertainty, but their standard errors are as well. 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 followed by the symbol "!". In such cases, the standard errors — and any confidence intervals or significance tests involving these standard errors — should be interpreted cautiously. Further details concerning procedures for identifying such standard errors are discussed in the *Technical Report of the NAEP 1996 State Assessment Program in Mathematics*.

Minimum Subgroup Sample Sizes

Results for mathematics performance and background variables were tabulated and reported for groups defined by gender, race/ethnicity, parental education, location of the school, type of school, participation in federally funded Title I programs, and eligibility for the free/reduced-price lunch component of the National School Lunch Program. NAEP collects data for five racial/ethnic subgroups (White, Black, Hispanic, Asian/Pacific Islander, and American Indian/Alaskan Native), three types of locations (Central City, Urban Fringe/Large Town, and Rural/Small Town),[10] and five levels of parents' education (Graduated From College, Some Education After High School, Graduated From High School, Did Not Finish High School, and I Don't Know).

[8] Benjamini, Y. and Y. Hochberg. "Controlling the false discovery rate: A practical and powerful approach to multiple testing," in *Journal of the Royal Statistical Society, Series B,* 57(1). (pp. 289—300, 1994).

[9] Williams, V.S.L., L.V. Jones, and J.W. Tukey. *Controlling Error in Multiple Comparisons, with Special Attention to the National Assessment of Educational Progress*. (Research Triangle Park, NC: National Institute of Statistical Sciences, December 1994).

[10] Previous NAEP reports reported data for four types of communities, rather than for the three types of location. These types of communities were Advantaged Urban, Disadvantaged Urban, Extreme Rural, and Other types of communities.

In many jurisdictions, and for some regions of the country, the number of students in some of these groups was not sufficiently high to permit accurate estimation of performance and/or background variable results. As a result, data are not provided for the subgroups with students from very few schools or for the subgroups with very small sample sizes. For results to be reported for any state assessment program subgroup, public school results must represent at least 5 primary sampling units (PSUs) and nonpublic school results must represent 6 schools. For results to be reported for any national assessment subgroup, at least 5 PSUs must be represented in the subgroup. In addition, a minimum sample of 62 students per subgroup is required. For statistical tests pertaining to subgroups, the sample size for both groups has to meet the minimum sample size requirements.

The minimum sample size of 62 was determined by computing the sample size required to detect an effect size of 0.5 total-group standard deviation units with a probability of 0.8 or greater. The effect size of 0.5 pertains to the *true* difference between the average scale score of the subgroup in question and the average scale score for the total fourth- or eighth-grade public school population in the jurisdiction, divided by the standard deviation of the scale score in the total population. If the *true* difference between subgroup and total group mean is 0.5 total-group standard deviation units, then a sample size of at least 62 is required to detect such a difference with a probability of 0.8. Further details about the procedure for determining minimum sample size appear in the *Technical Report of the NAEP 1996 State Assessment Program in Mathematics*.

Describing the Size of Percentages

Some of the percentages reported in the text of the report are given qualitative descriptions. For example, the number of students currently taking an algebra class might be described as "relatively few" or "almost all," depending on the size of the percentage in question. Any convention for choosing descriptive terms for the magnitude of percentages is to some degree arbitrary. The descriptive phrases used in the report and the rules used to select them are shown below.

Percentage	Descriptive Term Used in Report
$p = 0$	None
$0 < p \leq 8$	A small percentage
$8 < p \leq 13$	Relatively few
$13 < p \leq 18$	Less than one fifth
$18 < p \leq 22$	About one fifth
$22 < p \leq 27$	About one quarter
$27 < p \leq 30$	Less than one third
$30 < p \leq 36$	About one third
$36 < p \leq 47$	Less than half
$47 < p \leq 53$	About half
$53 < p \leq 64$	More than half
$64 < p \leq 71$	About two thirds
$71 < p \leq 79$	About three quarters
$79 < p \leq 89$	A large majority
$89 < p < 100$	Almost all
$p = 100$	All

A.4 Revisions to the NAEP 1990 and 1992 Mathematics Findings

After the NAEP 1994 assessment has been conducted, two technical problems were discovered in the procedures used to develop the NAEP mathematics scale and achievement levels determined for the 1990 and 1992 mathematics assessments. These errors affected the mathematics scale scores reported in 1992 and the achievement level results reported in 1990 mathematics scale scores reported in 1992 and the achievement level results reported in 1990 and 1992. The National Center for Education Statistics (NCES) and the National Assessment Governing Board (NAGB) have evaluated the impact of these errors and have reanalyzed and reported the revised results from both mathematics assessments. The technical errors have been corrected and the revised national and state scale score results for 1992 and achievement level results for 1990 and 1992 are presented in the NAEP 1996 mathematics reports.

Although the two technical problems that were discovered are discussed in greater detail in the *NAEP 1996 Technical Report* and *NAEP 1996 Technical Report of the State Assessment in Mathematics*, a brief summary is presented below.

The first technical problem resulted from an error in the computer program used to compute NAEP scale score results. The error occurred in the convention used to handle omitted responses in the item response theory (IRT) scaling of the partial-credit constructed-response questions, and it was limited only to those questions. In Analyses of the NAEP 1992 response questions, and it was limited only to those questions. In analyses of the NAEP 1992 mathematics assessment, this error caused all blank responses to partial-credit constructed-response questions (both omitted and not-reached responses) to be treated as missing — an acceptable treatment, but not the conventional choice for NAEP. (Because the NAEP 1990 mathematics assessment did not include these types of questions, the error did not occur.) The national and state assessments results were recalculated using the intended convention for the treatment of omitted responses.

In general, the effect of this technical problem on the previously reported NAEP 1992 mathematics findings was minimal, and it had little impact on policy-related interpretations. The recalculated 1992 mathematics scale score results, at the national and state levels, are quite similar to those published in the 1992 mathematics reports.

The second technical problem involved the development of the NAEP mathematics achievement level cut scores, and it concerned the mapping of the NAGB-approved achievement levels onto the NAEP mathematics scale. This error affected the achievement level results reported for the 1990 and 1992 mathematics assessments. In deriving the final levels recommended to NAGB, panelists' ratings for the multiple-choice and constructed-response questions were combined to obtain an overall rating for the questions. When combined, the ratings were weighted based on the amount of information provided by each type of question. In other words, some of the questions "counted more" toward the overall cut scores than others. However, because the weighting was carried out incorrectly, the constructed-response questions received more weight than intended. Therefore, the cut scores established by mapping the achievement levels onto the NAEP mathematics scale were incorrect, and the percentages of students at or above these levels were incorrectly estimated.

The program that mapped the achievement levels to the NAEP scale was corrected to appropriate weight the constructed-response questions, and revised mathematics achievement level cut scores were developed based on the corrected scaling procedures. As a result, the cut scores for the three achievement levels at each grade were raised, and the percentages of students at or above the achievement levels were recalculated based on the corrected cut scores. Revised 1990 and 1992 percentages, for the national and state assessments, are presented in this report.

APPENDIX B

The NAEP 1996 Mathematics Assessment

The 1996 assessment was the first update of the NAEP mathematics assessment framework[1] since the release of the National Council of Teachers of Mathematics (NCTM) *Curriculum and Evaluation Standards for School Mathematics*.[2] This update reflected refinements in the specifications governing the development of the 1996 assessment while assuring comparability of results across the 1990, 1992, and 1996 assessments. The refinements that distinguish the framework of the assessment conducted in 1996 from the framework of the assessments conducted in 1990 and 1992 include the following:

- moving away from the rigid content-strand-by-cognitive-process matrix that governed the development of earlier assessments. Classifying specific questions into cells of a matrix had required those questions to measure a unique content strand at a unique cognitive level. This stipulation often decontextualized the questions and limited the possibility of assessing students' abilities to reason in rich problem-solving situations and to make connections among content strands within mathematics.

- allowing individual questions on the assessment to be classified in one or more content strands when appropriate. Knowledge or skills from more than one content strand is often needed to answer a question. The option to classify questions in multiple ways provides a greater opportunity to measure student ability in content settings that closely approximate real-world reasoning and problem-solving situations. (However, to develop content strand scales, the primary content classification was used for questions with multiple classifications.)

- including the mathematics ability categories (conceptual understanding, procedural knowledge, and problem solving) as well as the process goals from the NCTM *Standards* (i.e., communication and connections) to achieve a balance of questions that measured a range of cognitive outcomes.

---

[1] National Assessment Governing Board. *Mathematics Framework for the 1996 National Assessment of Educational Progress*. (Washington, DC: National Assessment Governing Board, 1994).

[2] National Council of Teachers of Mathematics. *Curriculum and Evaluation Standards for School Mathematics*. (Reston, VA: NCTM, 1989).

- continuing the move towards including more constructed-response questions.

- creating "families" of questions that probe a student's understanding of mathematics vertically within a content strand or horizontally across content strands.

- revising the number sense, properties, and operations and geometry and spatial sense content strands to reflect the NCTM *Standards* emphasis on developing and assessing students' abilities to make sense of both number and operation and spatial settings.

These refinements to the NAEP mathematics framework were made so that the 1996 assessment would: (1) more adequately reflect recent curricular emphases and objectives and yet (2) maintain a connection with the 1990 and 1992 assessments to measure trends in student performance. Prior to the 1996 assessment, investigations were conducted to ensure that results from the assessment could be reported on the existing NAEP mathematics scale. The conclusion drawn from these investigations was that results from the 1990, 1992, and 1996 assessments could be reported on a common scale and trends in mathematics performance since 1990 examined.

The Assessment Design

Each student in the state assessment program in mathematics received a booklet containing a set of general background questions, a set of subject-specific background questions, and a combination of cognitive questions grouped in sets called blocks. At each grade level, the blocks of questions consisted of multiple-choice and constructed-response questions. Two types of constructed-response questions were included — short and extended constructed-response. Short constructed-response questions required students to provide answers to computation problems or to describe solutions in one or two sentences. Extended constructed-response questions required students to provide longer answers (e.g., a description of possibilities, a more involved computational analysis, or a description of a pattern and its implications). Students were expected to adequately answer the short constructed-response questions in about 2 to 3 minutes and the extended constructed-response questions in approximately 5 minutes. Short constructed-response questions which first appeared in the assessment in 1996 were graded to allow for partial credit (i.e., giving students credit for answers that are partially correct) according to a unique scoring rubric developed for each constructed-response question. Short constructed-response questions included in the 1990 and 1992 mathematics assessments were dichotomously scored (i.e., correct or incorrect). The extended constructed-response questions included in the 1992 and 1996 assessments were scored allowing for partial credit.

The blocks of questions contained several other features. Five to seven of the blocks at each grade level allowed calculator usage. At grade 4, students were provided four-function calculators, and at grade 8, students were provided scientific calculators. Prior to the assessment, all students were trained in the use of these calculators. For several blocks, students were given manipulatives (including geometric shapes, three-dimensional models, and spinners). For two of the blocks at each grade level, students were given rulers (at grade 4) or rulers and protractors (at grade 8) so the student could answer questions dealing with measurements and draw specified geometric shapes.

As part of the national assessment, other blocks of questions were developed for each of the grade levels. Each grade level had two estimation blocks that employed a paced-audiotape format to measure students' estimation skills. Each grade level also had two 30-minute theme blocks consisting of a mixture of multiple-choice and constructed-response questions. All of the questions in these blocks related to some aspect of a rich problem setting that served as a unifying theme for the entire block. Neither the estimation nor the theme block component were included in the state assessment program. Results for the estimation and theme blocks will be featured in future reports on the NAEP 1996 mathematics assessment.

Of the 17 blocks in the national sample at the fourth grade and the 19 blocks in the national sample at the eighth grade, 3 were carried forward from the 1990 assessment and 5 were carried forward from the 1992 assessment to allow for the measurement of trends across time. The remaining blocks of questions at each grade level contained new questions developed for the 1996 assessment as specified by the updated framework.

The data in Table B.1 reflect the number of questions by type by grade level for the 1990, 1992, and 1996 assessments. As mentioned earlier, the 1996 assessment continued NAEP's shift toward more constructed-response questions, including extended constructed-response questions that required students to provide an answer and a corresponding explanation.

TABLE B.1

Distribution of Questions by Question Type

	Grade 4			Grade 8			Grade 12		
	1990	1992	1996	1990	1992	1996	1990	1992	1996
Multiple-Choice	102	99	81	149	118	102	156	115	99
Short Constructed-Response*	41	59	64	42	65	69	47	64	74
Extended Constructed-Response**	---	5	13	---	6	12	---	6	11
Total	143	163	158	191	189	183	203	185	184

* Short constructed-response questions included in the 1990 and 1992 assessments were scored dichotomously. New short constructed-response questions included in the 1996 assessment were scored to allow for partial credit.
** No extended constructed-response questions were included in the 1990 assessment.

Each booklet in the state assessment program included three sets of student background questions. The first, consisting of general background questions, included questions about race or ethnicity, mother's and father's level of education, reading materials in the home, homework, attendance, and academic expectations. The second set, consisting of mathematics background questions, included questions about instructional activities, courses taken, use of specialized resources such as calculators in mathematics classes, and views on the utility and value of the subject. (Students were given 5 minutes to complete each set of questions, with the exception of the fourth graders, who were given more time because the general background questions were read aloud to them.) The third set of questions followed the cognitive question blocks and contained five questions about students' motivation to do well on the assessment, their perception of the difficulty of the assessment, and their familiarity with the types of cognitive questions included.

The blocks of cognitive and background questions were carefully balanced to ensure that the blocks could be completed within the time provided to the students, using information gathered from the field test. For more information on the design of the assessment, the reader is referred to Appendix C.

APPENDIX C

Technical Appendix: The Design, Implementation, and Analysis of the 1996 State Assessment Program in Mathematics

C.1 Overview

The purpose of this appendix is to provide technical information about the 1996 state assessment program in mathematics. It provides a description of the design for the assessment and gives an overview of the steps involved in the implementation of the program from the planning stages through to the analysis of the data.

This appendix is one of several documents that provide technical information about the 1996 state assessment program. Those interested in more details are referred to the forthcoming *Technical Report of the NAEP 1996 State Assessment Program in Mathematics*. Theoretical information about the models and procedures used in NAEP can be found in the special NAEP-related issue of the *Journal of Educational Statistics* (Summer 1992/Volume 17, Number 2) as well as previous national technical reports.

Educational Testing Service (ETS) was awarded the cooperative agreement for the 1996 NAEP programs, including the state assessment program. ETS was responsible for overall management of the programs as well as for development of the overall design, the cognitive questions and questionnaires, data analysis, and reporting. National Computer Systems (NCS) was a subcontractor to ETS on both the national and state NAEP programs. NCS was responsible for printing, distribution, and receipt of all assessment materials, and for scanning and professional scoring. All aspects of sampling and field operations for both the national and state assessment programs were the responsibility of Westat, Inc. NCES awarded a separate cooperative agreement to Westat for these services for the national and state assessments.

Organization of the Technical Appendix

This appendix provides a brief description of the design for the state assessment program in mathematics and gives an overview of the steps involved in implementing the program from the planning stages to the analysis of the data. (A more detailed discussion of the technical aspects of the NAEP state assessment program can be found in the forthcoming *Technical Report of the NAEP 1996 State Assessment Program in Mathematics*.) The organization of this appendix is as follows:

- Section C.2 provides an overview of the design of the 1996 state assessment program in mathematics.
- Section C.3 discusses the balanced incomplete block (BIB) spiral design that was used to assign cognitive questions to assessment booklets and assessment booklets to students.
- Section C.4 outlines the sampling design used for the 1996 state assessment program.
- Section C.5 summarizes Westat's field administration procedures.
- Section C.6 describes the flow of the data from their receipt at NCS through data entry and professional scoring.
- Section C.7 summarizes the procedures used to weight the assessment data and to obtain estimates of the sampling variability of subpopulation estimates.
- Section C.8 describes the initial analyses performed to verify the quality of the data.
- Section C.9 describes the item response theory scales and the overall mathematics composite scale that were created for the final analyses of the state assessment program data.
- Section C.10 provides an overview of the linking of the scaled results from the state assessment program in mathematics to those from the national assessment.

C.2 Design of the NAEP 1996 State Assessment Program in Mathematics

The major aspects of the design for the state assessment program in mathematics included the following:

- Participation at the jurisdiction level was voluntary.
- Fourth- and eighth-grade students from public and nonpublic schools were assessed. Nonpublic schools included Catholic schools, other religious schools, private schools, Department of Defense Domestic Elementary and Secondary Schools (DDESS), and Bureau of Indian Affairs schools. Separate representative samples of public and nonpublic schools were selected in each participating jurisdiction and students were randomly sampled within schools. The size of a jurisdiction's nonpublic school samples was proportional to the percentage of students in that jurisdiction attending such schools.

- The fourth- and eighth-grade mathematics assessment instruments used for the state assessment program and the national assessment consisted of 13 blocks of questions. Eight of these blocks were previously administered as part of the 1990 and 1992 national and Trial State Assessments. The type of questions — constructed-response or multiple-choice — was determined by the nature of the task. In addition, the constructed-response questions were of two types: *short constructed-response* questions required students to provide answers to computation problems or to describe solutions in one or two sentences, while *extended constructed-response* questions required students to provide longer responses when answering the question. Each student was given 3 of the 13 blocks of questions.

- A complex form of matrix sampling called a balanced incomplete block (BIB) spiraling design was used. With BIB spiraling, students in an assessment session received different booklets, which provided for greater mathematics content coverage than would have been possible had every student been administered the identical set of questions, without imposing an undue testing burden on the student.

- Background questionnaires given to the students, the students' mathematics teachers, and the principals or other administrators provided a variety of contextual information. The background questionnaires for the state assessment program were identical to those used in the national fourth- and eighth-grade assessments.

- The total assessment time for each student was approximately one hour and 40 minutes. Each assessed student was assigned a mathematics booklet that contained two 5-minute background questionnaires, followed by 3 of the 13 blocks of mathematics questions requiring 15 minutes each, and a 3-minute motivation questionnaire. Twenty-six different booklets were assembled.

- The assessments were scheduled to take place in the five-week period between January 29 and March 4, 1996. One-fourth of the schools in each jurisdiction were to be assessed each week throughout the first four weeks; however, due to the severe weather throughout much of the country, the fifth week was used for regular testing as well as for makeup sessions.

- Data collection was, by law, the responsibility of each participating jurisdiction. Security and uniform assessment administration were high priorities. Extensive training of state assessment personnel was conducted to assure that the assessment would be administered under standard, uniform procedures. For jurisdictions that had participated in previous NAEP state assessments, 25 percent of both public and nonpublic school assessment sessions were monitored by the Westat staff. For the jurisdictions new to NAEP, 50 percent of both public and nonpublic school sessions were monitored.

C.3 Assessment Instruments

The assembly of cognitive questions into booklets and their subsequent assignment to assessed students was determined by a BIB design with spiraled administration. This design is a variant of a matrix sampling design. The full set of mathematics questions was divided into 13 unique blocks, each requiring 15 minutes for completion. Each assessed student received a booklet containing 3 of the 13 blocks according to a design that ensured that each block was administered to a representative sample of students within each jurisdiction.

In addition to the student assessment booklets, three other instruments provided data relating to the assessment — a mathematics teacher questionnaire, a school characteristics and policies questionnaire, and an SD/LEP student questionnaire.

The *student assessment booklets* contained five sections and included both cognitive and noncognitive questions. In addition to three 15-minute sections of cognitive questions, each booklet included two 5-minute sets of general and mathematics background questions designed to gather contextual information about students, their experiences in mathematics, and their attitudes toward the subject, and one 3-minute section of motivation questions designed to gather information about the student's level of motivation while taking the assessment.

The *teacher questionnaire* was administered to the mathematics teachers of the fourth- and eighth-grade students participating in the assessment. The questionnaire consisted of three sections and took approximately 20 minutes to complete. The first section focused on the teacher's general background and experience; the second, on the teacher's background related to mathematics; and the third, on classroom information about mathematics instruction.

The *school characteristics and policies questionnaire* was given to the principal or other administrator in each participating school and took about 20 minutes to complete. The questions asked about the principal's background and experience, school policies, programs, and facilities, and the demographic composition and background of the students and teachers.

The *SD/LEP student questionnaire* was completed by the staff member most familiar with any student selected for the assessment who was classified in either of two ways: students with disabilities (SD) had an Individualized Education Plan (IEP) of equivalent special education plan (for reasons other than being gifted and talented); students with limited English proficiency were classified as LEP students. The questionnaire took approximately three minutes to complete and asked about the student and the special programs in which the student participated. It was completed for all selected SD or LEP students regardless of whether or not they participated in the assessment. Selected SD or LEP students participated in the assessment if they were determined by the school to be able to participate, considering the terms of their IEP and accommodations provided by the school or by NAEP.

C.4 The Sampling Design

The sampling design for NAEP is complex, in order to minimize burden on schools and students while maximizing the utility of the data; for further details see the forthcoming *Technical Report for the NAEP 1996 State Assessment Program in Mathematics*. The target populations for the state assessment program in mathematics consisted of fourth- and eighth-grade students enrolled in either public or nonpublic schools. The representative samples of public school fourth and eighth graders assessed in the state assessment program came from about 100 schools (per grade) in most jurisdictions. However, if a jurisdiction had fewer than 100 public schools with a particular grade, all or almost all schools were asked to participate. If a jurisdiction had smaller numbers of students in each school than expected, more than 100 schools were selected for participation. The nonpublic school samples differed in size across the jurisdictions, with the number of schools selected proportional to the nonpublic school enrollment within each jurisdiction. Typically, about 20 to 25 nonpublic schools (per grade) were included for each jurisdiction. The school sample in each jurisdiction was designed to produce aggregate estimates for the jurisdiction and for selected subpopulations (depending upon the size and distribution of the various subpopulations within the jurisdiction) and also to enable comparisons to be made, at the jurisdiction level, between administration of assessment tasks with monitoring and without monitoring. The public schools were stratified by urbanization, percentage of Black and Hispanic students enrolled, and median household income within the ZIP code area of the school. The nonpublic schools were stratified by type of control (Catholic, private/other religious, other nonpublic), metropolitan status, and enrollment size per grade.

The national and regional results presented in this report are based on nationally representative samples of fourth- and eighth-grade students. The samples were selected using a complex multistage sampling design involving the sampling of students from selected schools within selected geographic areas across the country. The sample design had the following stages:

(1) selection of geographic areas (a county, group of counties, or metropolitan statistical area)

(2) selection of schools (public and nonpublic) within the selected areas

(3) selection of students within selected schools

Each selected school that participated in the assessment, and each student assessed, represent a portion of the population of interest. To make valid inferences from student samples to the respective populations from which they were drawn, sampling weights are needed. Discussions of sampling weights and how they are used in analyses are presented in sections C.7 and C.8.

The state results provided in this report are based on state-level samples of fourth- and eighth-grade students. The samples of both public and nonpublic school students were selected based on a two-stage sample design that entailed selecting students within schools. The first-stage samples of schools were selected with a probability proportional to the fourth- or eighth-grade enrollment in the schools. Special procedures were used for jurisdictions with many small schools and for jurisdictions with a small number of schools. As with the national samples, the state samples were weighted to allow for valid inferences about the populations of interest.

The results presented for a particular jurisdiction are based on the representative sample of students who participated in the 1996 state assessment program. The results for the nation and regions of the country are based on the nationally and regionally representative samples of students who were assessed as part of the national NAEP program. Using the national and regional results from the 1996 national assessment was necessary because of the voluntary nature of the state assessment program. Because not every state participated in the program, the aggregated data across states did not necessarily provide representative national or regional results.

In most jurisdictions, up to 30 students were selected from each school, with the aim of providing an initial sample size of approximately 3,000 public school students per jurisdiction per grade. The student sample size of 30 for each school was chosen to ensure that at least 2,000 public school students (per grade) participated from each jurisdiction, allowing for school nonresponse, exclusion of students, inaccuracies in the measures of enrollment, and student absenteeism from the assessment. In jurisdictions with fewer schools, larger numbers of students per school were often required to ensure initial samples of roughly 3,000 students. In certain jurisdictions, all eligible fourth or eighth graders were targeted for assessment. Jurisdictions were given the option to reduce the expected student sample size in order to reduce testing burden and the number of multiple-testing sessions for participating schools. At grade 4, two jurisdictions (Delaware and Guam) and at grade 8, four jurisdictions (Alaska, Delaware, Hawaii, and Rhode Island) elected to exercise this option. Using this option can involve compromises such as higher standard errors and accompanying loss of precision.

In order to provide for wider inclusion of students with disabilities and limited English proficiency, the 1996 state assessments in mathematics involved dividing the sample of students at each grade level into two subsamples, referred to as S1 and S2. S1 provided continuity with the 1992 mathematics assessment and thus allowed for the reporting of performance over time by using the same exclusion criteria for students with disabilities and limited English proficiency as was used in that assessment. S2 provided for wider inclusion of students with disabilities and limited English proficiency by incorporating new exclusion rules. For further discussion, see the *NAEP 1996 Mathematics Report Card*. The 1996 national assessment in mathematics involved an additional subsample, S3, in which accommodations were provided for certain students with disabilities or limited English proficiency, again in order to make NAEP more inclusive.

For both the national and state mathematics assessments, scaling and analysis procedures (discussed in sections C.8 to C.10) were applied to a combination of students from S1 and S2. Specifically, all assessed students from S1 were combined with those students from S2 who were **not** identified as SD or LEP. This combination of segments of the S1 and S2 subsamples provided for maximizing the use of available data while allowing for comparisons to the student population in the national sample. This combination, referred to as the "reporting sample," was the sample used in linking the state assessment to the national assessment (see Section C.10).

Additional analyses will be conducted on the national samples in order to study the effects of changing the exclusion rules and the presence of accommodations. Preliminary discussion can be found in the *NAEP 1996 Mathematics Report Card* and more detailed discussion will follow in future NAEP publications.

C.5 Field Administration

The administration of the 1996 program required collaboration between staff in the participating jurisdictions and schools and the NAEP contractors, especially Westat, the field administration contractor.

Each jurisdiction volunteering to participate in the 1996 state assessment program was asked to appoint a state coordinator as liaison between NAEP staff and the participating schools. In addition, Westat hired and trained a supervisor for each jurisdiction and six field managers, each of whom was assigned to work with groups of jurisdictions. The state supervisors were responsible for working with the state coordinators, overseeing assessment activities, training school district personnel to administer the assessment, and coordinating the quality-control monitoring efforts. Each field manager was responsible for working with the state coordinators of seven to eight jurisdictions and for the supervision of the state supervisors assigned to those jurisdictions. An assessment administrator was responsible for preparing for and conducting the assessment session in one or more schools. These individuals were usually school or district staff and were trained by Westat. Westat also hired and trained three to five quality control monitors in each jurisdiction. For jurisdictions that had previously participated in the state assessment program, 25 percent of the public and nonpublic school sessions were monitored. For jurisdictions new to the program, 50 percent of all sessions were monitored. The assessment sessions were conducted during a five-week period beginning in late January 1996.

C.6 Materials Processing, Professional Scoring, and Database Creation

Upon completion of each assessment session, school personnel shipped the assessment booklets and forms to NCS for professional scoring, entry into computer files, and checking. The files were then sent to ETS for creation of the database.

After NCS received all appropriate materials from a school, they were forwarded to the professional scoring area where the responses to the constructed-response question were evaluated by trained staff using guidelines prepared by ETS. Each constructed-response question had a unique scoring guide that defined the criteria to be used in evaluating students' responses. The extended constructed-response questions were evaluated with four- or five-level rubrics, and the short constructed-response questions first used in 1996 were rated according to three-level rubrics that permit partial credit to be given. Short constructed-response questions used previously were scored dichotomously (i.e., correct or incorrect).

For the national mathematics assessment and the state assessment program in mathematics, over 4.8 million constructed responses were scored. This figure includes rescoring to monitor inter-rater reliability and trend reliability. In other words, scoring reliability was calculated both within year (1996) and across years (1990, 1992, and 1996). The overall within-year percentages of agreement for the 1996 national within-year reliability samples were 96 percent at grade 4 and 96 percent at grade 8. The percentages of agreement across the assessment years for the national inter-year reliability samples were 96 percent (1990 to 1996) and 94 percent (1992 to 1996) at grade 4 and 95 percent (1990 to 1996) and 94 percent (1992 to 1996) at grade 8.

Data transcription and editing procedures were used to generate the disk and tape files containing various assessment information, including the sampling weights required to make valid statistical inferences about the population from which the state assessment program sample was drawn. Prior to analysis, the data from these files underwent a quality control check at ETS. The files were then merged into a comprehensive, integrated database.

C.7 Weighting and Variance Estimation

A complex sample design was used to select the students to be assessed in each of the participating jurisdictions. The properties of a sample from a complex design are very different from those of a simple random sample in which every student in the target population has an equal chance of selection and in which the observations from different sampled students can be considered to be statistically independent of one another. The properties of the sample from the complex state assessment program design were taken into account in the analysis of the assessment data.

One way that the properties of the sample design were addressed was by using sampling weights to account for the fact that the probabilities of selection were not identical for all students. These weights also included adjustments for school and student nonresponse. All population and subpopulation characteristics based on the state assessment program data used sampling weights in their estimation.

In addition to deriving appropriate estimates of population characteristics, it is essential to obtain appropriate measures of the degree of uncertainty of those statistics. One component of uncertainty results from sampling variability, which is a measure of the dependence of the results on the particular sample of students actually assessed. Because of the effects of cluster selection (schools are selected first, then students are selected within those schools), observations made on different students cannot be assumed to be independent of each other (and, in fact, are generally positively correlated). As a result, classical variance estimation formulas will produce incorrect results. Instead, a jackknife variance estimation procedure that takes the characteristics of the sample into account was used for all analyses.

Jackknife variance estimation provides a reasonable measure of uncertainty for any statistic based on values observed without error. Statistics such as the percentage of students correctly answering a given question meet this requirement, but other statistics based on estimates of student mathematics performance, such as the average mathematics scale score of a subpopulation, do not. Because each student typically responds to relatively few questions from a particular content strand (e.g., Algebra and Functions or Geometry and Spatial Sense) there exists a nontrivial amount of imprecision in the measurement of the scale score of a given student. This imprecision adds an additional component of variability to statistics based on estimates of individual scale scores.

C.8 Preliminary Data Analysis

After the computer files of student responses were received from NCS and merged into an integrated database, all cognitive and noncognitive questions were subjected to an extensive item analysis. For each question, this analysis yielded the number of respondents, the percentage of responses in each category, the percentage who omitted the question, the percentage who did not reach the question, and the correlation between the question score and the block score. In addition, the item analysis program provided summary statistics for each block, including a reliability (internal consistency) coefficient. These analyses were used to check the scoring of the questions, to verify the appropriateness of the difficulty level of the questions, and to check for speededness. The results were reviewed by knowledgeable project staff in search of aberrations that might signal unusual results or errors in the database.

The question and block-level analyses were done using rescaled versions of the final sampling weights provided by Westat (see Section C.7). The rescaling was carried out within each jurisdiction. The sum of the sampling weights for the public school students within each jurisdiction was constrained to be equal. The same transformation was then applied to the weights of the nonpublic school students in that jurisdiction. The sum of the weights for each of the DoDEA samples (i.e., DDESS and DoDDS) was constrained to be equal to the same value as the public school students in other jurisdictions. Use of rescaled weights does nothing to alter the value of statistics calculated separately within each jurisdiction. However, for statistics obtained from samples that combine students from different jurisdictions, use of the rescaled weights results in a roughly equal contribution of each jurisdiction's data to the final value of the estimate. Equal contribution of each jurisdiction's data to the results of the item response theory (IRT) scaling was viewed as a desirable outcome. The original final sampling weights provided by Westat were used in reporting.

Additional analyses comparing the data from the monitored sessions with those from the unmonitored sessions were conducted to determine the comparability of the assessment data from the two types of administrations. Differential item functioning (DIF) analyses were carried out using the national assessment data. DIF analyses identify questions that were differentially difficult for various subgroups, affording the opportunity to reexamine such questions with respect to their fairness and their appropriateness for inclusion in the scaling process.

C.9 Scaling the Assessment Questions

The primary analysis and reporting of the results from the state assessment program used item response theory (IRT) scale-score models. Scaling models quantify a respondent's tendency to provide correct answers to the domain of questions contributing to a scale as a function of a parameter called performance, estimated by a scale score. The scale scores can be viewed as a summary measure of performance across the domain of questions that make up the scale. Three distinct IRT models were used for scaling: 1) 3-parameter logistic models for multiple-choice questions; 2) 2-parameter logistic models for short constructed-response questions that were scored correct or incorrect; and 3) generalized partial credit models for short and extended constructed-response questions that were scored on a multipoint (i.e., greater than two levels) scale.

Five distinct scales were created for the state assessment program in mathematics to summarize fourth- and eighth-grade students' abilities according to the five defined content strands (Number Sense, Properties, and Operations; Measurement; Geometry and Spatial Sense; Data Analysis, Statistics, and Probability; and Algebra and Functions). These scales were defined identically to, but separately from, those used for the scaling of the national NAEP fourth- and eighth-grade mathematics data. Although the questions comprising each scale were identical to those used in the national assessment program, the item parameters for the state assessment program scales were estimated from combined public school data from the jurisdictions participating in the state assessment program.[1] Item parameter estimation was carried out on an item calibration subsample. The calibration subsample consisted of an approximately 25 percent sample of all available public school data. To ensure equal representation in the scaling process, each jurisdiction contributed the same number of students to the item calibration sample. Within each jurisdiction, 50 percent of the calibration sample was taken from monitored administrations and the other 50 percent came from unmonitored administrations.

The fit of the IRT model to the observed data was examined within each scale by comparing the estimates of the empirical item characteristic functions with the theoretic curves. For correct-incorrect questions, nonmodel-based estimates of the expected proportions of correct responses to each question for students with various levels of scale proficiency were compared with the fitted item response curve; for the short and extended partial-credit constructed-response questions, the comparisons were based on the expected proportions of students with various levels of scale proficiency who achieved each score level. In general, the question-level results were well fit by the scaling models.

[1] Schools from the DoDEA jurisdictions were not included in the item calibration sample.

Using the item parameter estimates, estimates of various population statistics were obtained for each jurisdiction. The NAEP methods use random draws ("plausible values") from estimated proficiency distributions for each student to compute population statistics. Plausible values are not optimal estimates of individual student proficiencies; instead, they serve as intermediate values to be used in estimating population characteristics. Under the assumptions of the scaling models, these population estimates will be consistent, in the sense that the estimates approach the model-based population values as the sample size increases, which would not be the case for population estimates obtained by aggregating optimal estimates of individual performance.

In addition to the plausible values for each scale, a composite of the five content strand scales was created as a measure of overall mathematics proficiency. This composite was a weighted average of the five mathematics scales in which the weights were proportional to the relative importance assigned to each content strand in the mathematics framework. The definition of the composite for the state assessment program was identical to that used for the national fourth- and eighth-grade mathematics assessments.

C.10 Linking the State Results to the National Results

A major purpose of the state assessment program was to allow each participating jurisdiction to compare its 1996 results with those for the nation as a whole and with those for the region of the country in which that jurisdiction is located. For meaningful comparisons to be made between each jurisdiction and the relevant national sample, results from these two assessments had to be expressed in terms of a similar system of scale units.

The results from the state assessment program were linked to those from the national assessment through linking functions determined by comparing the results for the aggregate of all students assessed in the state assessment program with the results for students of the matching grade within the National Linking Sample of the national NAEP. The National Linking Sample of the national NAEP for a given grade is a representative sample of the population of all grade-eligible public school students within the aggregate of 45 participating states and the District of Columbia. Guam and the two Department of Defense Education Activity (DoDEA) jurisdictions were not included in the aggregate. Specifically, the fourth- and eighth-grade National Linking Samples consist of all fourth- and eighth-grade students in public schools in the states and the District of Columbia who were assessed in the national cross-sectional mathematics assessment.

For each grade, a linear equating within each scale was used to link the results of the state assessment program to the national assessment. For each scale, the adequacy of the linear equating was evaluated by comparing the distribution of mathematics scale scores based on the aggregation of all assessed students at each grade from the participating states and the District of Columbia with the equivalent distribution based on the students in the National Linking Sample. In the estimation of these distributions, the students were weighted to represent the target population of public school students in the specified grade in the aggregation of the states and the District of Columbia. If a linear equating were adequate, the distribution for the aggregate of states and the District of Columbia and that for the National Linking Sample will have, to a close approximation, the same shape in terms of the skewness, kurtosis, and higher moments of the distributions. The only differences in the distributions allowed by linear equating are in the means and variances. Generally, this has been found to be the case.

Each mathematics content-strand scale was linked by matching the mean and standard deviation of the scale scores across all students in the state assessment (excluding Guam and the two DoDEA jurisdictions) to the corresponding scale mean and standard deviation across all students in the National Linking Sample.

APPENDIX D

Setting the Achievement Levels

Setting achievement levels is a test-centered method for setting standards on the NAEP assessment that identifies what students should know and should be able to do. The method depends on securing and summarizing a set of judgmental ratings of expectations for student educational performance on specific questions comprising the NAEP mathematics assessment. The NAEP mathematics scale is a numerical index of students' performance in mathematics ranging from 0 to 500. The three achievement levels — *Basic, Proficient,* and *Advanced* — are mapped onto the scale for each grade level assessed.

The NAEP mathematics achievement levels were set following the 1990 assessment and further refined following the 1992 assessment. In developing the threshold values for the levels, a broadly constituted panel of judges — including teachers (50%), non-teacher educators (20%), and the general public (noneducators)[1] (30%) — rated a grade-specific item pool using the policy definitions of the National Assessment Governing Board (NAGB) for *Basic, Proficient*, and *Advanced*. The policy definitions were operationalized by the judges in terms of specific mathematical skills, knowledge, and behaviors that were judged to be appropriate expectations for students in each grade and were in accordance with the current mathematics assessment framework. The policy definitions are as follows:

Basic

This level denotes partial mastery of the prerequisite knowledge and skills that are fundamental for proficient work at each grade.

Proficient

This level represents solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter and are well prepared for the next level of schooling.

Advanced

This higher level signifies superior performance beyond proficient grade-level mastery at each grade.

[1] Noneducators represented business, labor, government service, parents, and the general public.

The judges' operationalized definitions were incorporated into lists of descriptors that represent what borderline students should be able to do at each of the levels defined by policy. The purpose of having panelists develop their own operational definitions of the achievement levels was to ensure that all panelists would have a common understanding of borderline performances and a common set of content-based referents to use during the item-rating process.

The judges (24 at grade 4 and 22 at grade 8) each rated half of the questions in the NAEP pool in terms of the expected probability that a student at a borderline achievement level would answer the question correctly, based on the judges' operationalization of the policy definitions and the factors that influence question difficulty. To assist the judges in generating consistently scaled ratings, the rating process was repeated twice, with feedback. Information on consistency among different judges and on the difficulty of each question[2] was fed back into the first repetition (round 2), while information on consistency within each judge's set of ratings was fed back into the second repetition (round 3). The third round of ratings permitted the judges to discuss their ratings among themselves to resolve problematic ratings. The mean final rating of the judges aggregated across questions yielded the threshold values in the percent correct metric. These cut scores were then mapped onto the NAEP scale (which is defined and scored using item response theory, rather than percent correct) to obtain the scale scores for the achievement levels.[3] The judges' ratings, in both metrics, and their associated errors of measurement are shown below. NAGB accepted the panel's achievement levels and, for reporting purposes, set final cutpoints one standard error (a measure of consistency among the judges' ratings) below the mean levels.

THE NATION'S REPORT CARD naep 1996 State Assessment

FIGURE D.1

Cutpoints for Achievement Levels at Grades 4 and 8

Grade	Level	Mean Percent Correct (Round 3)	Scale Score*	Standard Error of Scale Score**
4	Basic	39	214	1.9
4	Proficient	65	249	4.1
4	Advanced	84	282	4.0
8	Basic	48	262	2.4
8	Proficient	71	299	5.7
8	Advanced	87	333	4.8

* Scale score is derived from a weighted average of the mean percent correct for multiple-choice and short constructed-response questions after both were mapped onto the NAEP scale.

** The standard error of the scale score is estimated from the difference in mean scale scores for the two equivalent subgroups of judges.

[2] Item difficulty estimates were based on a preliminary, partial set of responses to the national assessment.

[3] See Appendix A for a discussion of the technical errors that resulted in the reanalysis and rereporting of 1990 and 1992 mathematics achievement level results.

After the ratings were completed, the judges for each grade level reviewed the operationalized descriptions developed by the judges of the other grade levels as well as their own descriptions and defined achievement level descriptions that were generally acceptable to all three grade-group judges. However, the descriptions varied in format, sharpness of language, and degree of specificity of the statements. Therefore, another panel at a subsequent validation meeting improved the wording and modified the language of the achievement level descriptions to reflect more closely the terminology of the National Council of Teachers of Mathematics *Curriculum and Evaluation Standards for School Mathematics*.[4] The achievement level descriptions, though based on the 1992 NAEP pool, apply to the current assessment and will not change from assessment to assessment (that is, until the framework changes).

Figure 3.1 in Chapter 3 provides the detailed descriptions of the three achievement levels for grade 4. In addition, exemplar questions are presented to illustrate each level.

[4] National Council of Teachers of Mathematics. *Curriculum and Evaluation Standards for School Mathematics.* (Reston, VA: NCTM, 1989).

APPENDIX E

Teacher Preparation

Teachers are key to improving mathematics learning, and so it is important to examine their background and professional development. Fourth-grade mathematics teachers completed questionnaires concerning their background and training, including their experience, certification, undergraduate and graduate course work in mathematics, and involvement in pre-service education.

Consistent with procedures used throughout this report, the student was the unit of analysis. That is, the mathematics teachers' responses were linked to their students, and the data reported are the percentages of students taught by teachers with particular characteristics.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE E.1 — GRADE 4

Public School Teachers' Reports on Their Highest Level of Education

What is the highest academic degree you hold?		New Jersey	Northeast	Nation
		Percentage of Students		
Bachelor's degree	1992	66 (3.9)	40 (6.1)	53 (2.4)
	1996	68 (3.0)	45 (7.6)	60 (2.8)
Master's degree	1992	28 (3.8)	48 (5.4)	41 (2.4)
	1996	27 (2.3)	45 (6.6)	32 (2.3)<
Education specialist's or professional diploma	1992	6 (1.8)	10 (2.7)	6 (1.2)
	1996	5 (1.8)	10 (2.0)	8 (1.2)
Doctorate or professional degree	1992	1 (****)	2 (****)	0 (0.3)
	1996	0 (****)	0 (****)	0 (0.0)

The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. **** Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE E.2 — GRADE 4

Public School Teachers' Reports on Their Undergraduate Majors

What were your undergraduate major fields of study?	New Jersey	Northeast	Nation
	Percentage of Students		
GRADE 4			
Education	91 (1.8)	87 (4.0)	88 (1.5)
Mathematics	2 (0.8)	10 (3.8)	7 (1.4)
Mathematics Education	1 (0.6)	7 (3.2)	6 (1.2)
Special Education	4 (1.8)	4 (2.0)	8 (1.5)
ESL	0 (****)	2 (****)	3 (0.9)
Other	32 (2.7)	36 (5.1)	37 (2.5)

The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). **** Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

TABLE E.3 — GRADE 4

Public School Teachers' Reports on Their Graduate Majors

What were your graduate major fields of study?	New Jersey	Northeast	Nation
	Percentage of Students		
GRADE 4			
Education	34 (3.4)	69 (5.4)	59 (2.7)
Mathematics	1 (0.7)	2 (0.7)	4 (1.0)
Mathematics Education	2 (0.7)	4 (1.6)	5 (1.3)
Special Education	5 (1.7)	9 (4.0)	5 (1.2)
Bilingual	1 (0.5)	0 (****)	2 (0.6)
Admin./Supervision/Curric.	12 (2.4)	11 (3.0)	15 (1.8)
Counseling	4 (1.2)	2 (****)	2 (0.7)
Other	20 (2.7)	18 (2.2)	16 (1.7)
No graduate study	46 (3.3)	24 (5.0)	31 (2.4)

The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). **** Standard error estimates cannot be accurately determined.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

TABLE E.4 — GRADE 4

Public School Teachers' Reports on Their Teaching Certification

What type of teaching certification do you have in this state in your main assignment field?	New Jersey	Northeast	Nation
	Percentage of Students		
GRADE 4			
None, Accreditation other than state, Temporary, or Probationary	4 (1.3)	11 (3.4)	5 (1.0)
Regular	88 (2.8)	71 (3.1)	79 (2.0)
Advanced	8 (2.0)	18 (2.1)	16 (1.8)

The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details).
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

TABLE E.5 — GRADE 4

Public School Teachers' Reports on Years Teaching Experience

Counting this year, how many years in total have you taught . . .		New Jersey	Northeast	Nation
		Percentage of Students		
At either the elementary or secondary level				
2 years or less	1992	5 (1.1)	10 (3.8)	8 (1.1)
	1996	6 (1.5)	11 (3.5)	8 (1.3)
3-5 years	1992	5 (1.6)	8 (2.9)	13 (1.7)
	1996	8 (1.6)	10 (3.4)	14 (1.6)
6-10 years	1992	11 (1.7)	7 (2.2)	14 (1.6)
	1996	12 (2.7)	17 (4.6)	23 (2.0)>
11-24 years	1992	55 (3.3)	58 (5.1)	46 (2.4)
	1996	44 (3.6)	30 (6.0)<	35 (2.6)<
25 years or more	1992	24 (2.9)	17 (2.9)	19 (1.6)
	1996	29 (3.5)	32 (6.6)	20 (2.5)
Mathematics*				
2 years or less	1996	7 (1.6)	13 (4.0)	11 (1.6)
3-5 years	1996	9 (1.6)	11 (3.8)	14 (1.9)
6-10 years	1996	12 (2.6)	15 (4.4)	26 (2.0)
11-24 years	1996	45 (3.7)	33 (7.1)	33 (2.8)
25 years or more	1996	27 (3.4)	28 (6.9)	16 (2.1)

The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level. * This question was not asked of fourth-grade teachers in 1992.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

TABLE E.6 — GRADE 4

Public School Teachers' Reports on Courses in Mathematics or Mathematics Education

During the last two years, how many college or university courses have you taken in mathematics or mathematics education?	New Jersey	Northeast	Nation
	Percentage of Students		
GRADE 4			
None	89 (2.1)	84 (2.9)	78 (2.4)
One	6 (1.5)	12 (2.1)	14 (1.8)
Two	2 (1.0)	3 (1.0)	5 (1.1)
Three or more	3 (1.2)	1 (0.7)	4 (1.2)

The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details).
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

TABLE E.7 — GRADE 4

Public School Teachers' Reports on Coursework in the Use of Technology

During the past five years, have you taken courses or participated in professional development activities in the use of technology such as computers?	New Jersey	Northeast	Nation
	Percentage of Students		
GRADE 4			
Yes	70 (3.1)	77 (3.4)	83 (1.8)

The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details).
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

TABLE E.8 — GRADE 4

Public School Teachers' Reports on Studies of Mathematics Instruction Techniques

Have you ever studied any of the following, either in college or university courses or in professional development workshops or seminars?		New Jersey	Northeast	Nation
		Percentage of Students		
Estimation	1992	80 (2.8)	76 (4.2)	82 (1.9)
	1996	76 (2.8)	65 (3.9)	77 (1.7)
Problem solving in mathematics	1992	95 (1.4)	87 (4.0)	92 (1.3)
	1996	91 (2.2)	76 (5.4)	90 (1.6)
Use of manipulatives in mathematics instruction	1992	92 (2.2)	91 (3.4)	93 (1.0)
	1996	93 (1.6)	78 (3.9)	91 (1.2)
Use of calculators in mathematics instruction	1992	57 (2.9)	61 (4.1)	60 (2.7)
	1996	78 (3.1)>	56 (3.7)	72 (2.1)>
Understanding students' thinking about math	1992	74 (2.9)	72 (5.1)	72 (2.1)
	1996	68 (3.1)	62 (4.1)	71 (1.9)

The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

TABLE E.9 — GRADE 4

Public School Teachers' Reports on Studies of Gender and Cultural Issues

Have you ever studied any of the following, either in college or university courses or in professional development workshops or seminars?		New Jersey	Northeast	Nation
		Percentage of Students		
Gender issues in the teaching of mathematics	1992	38 (3.1)	30 (5.4)	33 (2.4)
	1996	41 (4.1)	31 (5.3)	46 (2.5)>
Teaching students from different cultural backgrounds	1992	48 (3.6)	38 (5.5)	43 (2.5)
	1996	43 (3.8)	39 (5.7)	50 (2.7)

The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details). If the notation > (<) appears, it signifies that the value for public school students was significantly higher (lower) than the value for 1992 at about the 95 percent confidence level.
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992 and 1996 Mathematics Assessments.

THE NATION'S REPORT CARD naep 1996 State Assessment

TABLE E.10 — GRADE 4

Public School Teachers' Reports on Professional Development

During the last year, how much time in total have you spent in professional development workshops or seminars in mathematics or mathematics education?	New Jersey	Northeast	Nation
	Percentage of Students		
GRADE 4			
None	14 (2.2)	19 (6.2)	15 (2.1)
Less than 6 hours	38 (4.2)	23 (3.9)	29 (2.2)
6-15 hours	26 (3.6)	30 (5.9)	28 (2.4)
16-35 hours	15 (3.2)	21 (5.7)	15 (2.0)
More than 35 hours	7 (1.6)	7 (1.1)	13 (1.6)

The standard errors of the statistics appear in parentheses. It can be said with about 95 percent confidence that, for each population of interest, the value for the entire population is within ± 2 standard errors of the estimate for the sample. In comparing two estimates, one must use the standard error of the difference (see Appendix A for details).
SOURCE: National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996 Mathematics Assessment.

ACKNOWLEDGMENTS

This report is the culmination of the efforts of many individuals who contributed their considerable knowledge, experience, and creativity to the NAEP 1996 mathematics assessment. The NAEP 1996 mathematics state assessment was a collaborative effort among staff from the National Center for Education Statistics (NCES), the National Assessment Governing Board (NAGB), Educational Testing Service (ETS), Westat, Inc., and National Computer Systems (NCS). In addition, 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 provided their wisdom, experience, and hard work. Most importantly, NAEP is grateful to the over 239,000 students and the teachers and administrators in over 9,700 schools in 48 jurisdictions who made the assessment possible by contributing considerable amounts of time and effort.

The NAEP 1996 mathematics state assessment was funded through NCES, in the Office of Educational Research and Improvement of the U.S. Department of Education. The Commissioner of Education Statistics, Pascal D. Forgione, Jr., and the NCES staff — Sue Ahmed, Peggy Carr, Arnold Goldstein, Steven Gorman, Larry Ogle, Gary W. Phillips, Sharif Shakrani, Maureen Treacy — and Alan Vanneman of the Education Statistics Services Institute, worked closely and collegially with the authors to produce this report. The authors were also provided invaluable advice and guidance by the members of the National Assessment Governing Board and NAGB staff. In particular, the authors are indebted to Arnold Goldstein of NCES for his daily efforts to coordinate the activities of the many people who contributed to this report.

The NAEP project at ETS is housed in the Center for the Assessment of Educational Progress under the direction of Paul Williams. The NAEP 1996 assessments were directed by Stephen Lazer and John Mazzeo. Jeff Haberstroh directed the scoring operations for the 1996 mathematics assessment. Sampling and data collection activities were conducted by Westat under the direction of Rene Slobasky, Nancy Caldwell, Keith Rust, Debby Vivari, and Dianne Walsh. Printing, distribution, scoring, and processing activities were conducted by NCS under the direction of Brad Thayer, Patrick Bourgeacq, Charles Brungardt, Mathilde Kennel, Linda Reynolds, and Connie Smith.

The complex statistical and psychometric activities necessary to report results for the NAEP 1996 mathematics assessment were directed by Nancy Allen, John Barone, James Carlson, and Juliet Shaffer. John Mazzeo and Gene Johnson provided direction on several critical psychometric issues. The analyses presented in this report were led by Frank Jenkins and Edward Kulick, with assistance from Hua Chang, Steve Wang, Xiaohui Wang, Hong Zhou, Jiahe Qian, Kate Pashley, David Freund, and Norma Norris.

Laura Jerry was responsible for the development and creation of the computer-generated reports, with assistance from Xiaohui Wang, Laura Jenkins, Phillip Leung, Inge Novatkoski, Bruce Kaplan, and Alfred Rogers. A large group of NAEP staff at ETS checked the data, text, and tables. Debbie Kline coordinated the technical appendices.

Many thanks are due to the comments and critical feedback of numerous reviewers, both internal and external to NCES and ETS. Important contributions were made by reviewers from academic institutions and education agencies: Bruce Brombacher of Upper Arlington (Ohio) Schools; Pasquale DeVito of the Rhode Island Department of Education, John Dossey of Illinois State University, Thomas Fisher of the Florida Department of Education, Douglas Rindone of the Connecticut Department of Education, and Irvin Vance of Michigan State University. Valuable input was given by NAGB staff Mary Lynn Bourque and Lawrence Feinberg, and NCES staff Susan Ahmed, Peggy Carr (who helped guide the report through several versions), Steven Gorman, Andrew Kolstad, Mary Frase, Mary Rollefson, Sharif Shakrani, and Shi-Chang Wu.

Cover design and production of the print version was directed by Carol Errickson, with the assistance of Sharon Davis-Johnson, Alice Kass, and Barbette Tardugno. Karen Damiano produced tables and text for one state for which the computerized report generating system was not appropriate. The World Wide Web version of the state reports was produced by Phillip Leung and Pat O'Reilly with assistance from Debbie Kline, Karen Damiano, Sharon M. Davis-Johnson, Craig Pizzuti, Barbette Tardugno, and Christine Zelenak. The *NAEP 1996 Mathematics State Report* for all participating jurisdictions is available via **http://www.ed.gov/NCES/naep**.

NCES 97-974