1999 Texas National Comparative Data Study

1999 Texas National Comparative Data Study

Executive Summary

Texas Education Code Section 35.028 requires that the state assessment program obtain national comparative results for the subject areas and grade levels for which criterion-referenced assessment instruments (TAAS) are administered. The Texas Education Agency contracted with Harcourt Educational Measurement (HEM) to conduct a study to obtain the required national comparative data. A stratified cluster sample of non-special education and non-gifted/talented classrooms from Texas public schools was taken in the fall of 1998 from each of grades 3–8 and 10; students in these classes were administered the *Metropolitan Achievements Tests*, Seventh Edition (MAT-7) in the spring of 1999 to measure student performance at the state level. The mathematics and reading subtests of MAT-7 were administered to students in all sampled grades, the language (similar to the multiple-choice writing TAAS) subtest of MAT-7 was administered to students in grades 4, 8, and 10, and the science and social studies subtests of MAT-7 were administered to students in grade 8.

Results were reported in terms of mean national normal curve equivalents (NCEs) based on the 1997 standardization of MAT-7. The mean national NCE obtained by the Texas student sample was above the national norm sample mean on the mathematics, language, science and social studies subtests for all grades. On the reading subtest, the mean national NCE obtained by the Texas student sample was slightly higher than the national norm sample mean in grades 3–5 and 10 and slightly lower than the national norm sample mean in grades 6–8.

Mean National NCEs and Associated Standard Errors of the Mean

	Grade						
	3	4	5	6	7	8	10
Total Reading							
Mean National NCE	50.3	52.0	51.2	48.9	49.9	48.4	51.7
Standard Error of the Mean	(0.39)	(0.45)	(0.45)	(0.63)	(0.65)	(0.66)	(0.80)
Total Mathematics							
Mean National NCE	55.4	59.4	57.1	54.8	55.1	53.4	54.9
Standard Error of the Mean	(0.46)	(0.40)	(0.46)	(0.58)	(0.67)	(0.64)	(0.70)
Language							
Mean National NCE		53.8				50.6	51.6
Standard Error of the Mean		(0.41)				(0.71)	(0.66)
Science							
Mean National NCE						52.5	
Standard Error of the Mean						(0.52)	
Social Studies							
Mean National NCE						50.5	
Standard Error of the Mean						(0.50)	

Introduction

Texas Education Code Section 35.028 requires that the state assessment program obtain national comparative results for the subject areas and grade levels for which criterion-referenced assessment instruments (TAAS) are administered. The Texas Education Agency has contracted with Harcourt Educational Measurement (HEM) to conduct a study to obtain the required national comparative data. A stratified cluster sample of non-special education and non-gifted/talented classrooms from Texas public schools was taken in the fall of 1998 from each of grades 3–8 and 10; students in these classes were administered the *Metropolitan Achievements Tests*, Seventh Edition (MAT-7) in the spring of 1999 to measure student performance at the state level.

Students were tested in grades 3, 4, 5, 6, 7, 8, and 10. The reading, mathematics, language (similar to the multiple-choice writing TAAS), science, and social studies subtests of MAT-7 were administered. The table below shows which grade levels and content areas were tested, as well as the approximate number of students selected in the target sample.

MAT-7 Target Sample

Content Area						
Grade	Reading	Math	Language	Science	Social Studies	Totals
3	12,500	12,500				25,000
4	12,500	12,500	12,500			37,500
5	12,500	12,500				25,000
6	12,500	12,500				25,000
7	12,500	12,500				25,000
8	12,500	12,500	12,500	12,500	12,500	62,500
10	12,500	12,500	12,500			37,500
Totals	87,500	87,500	37,500	12,500	12,500	237,500

The national norms used for the Texas National Comparative Data Study were obtained from the 1997 standardization of MAT-7. The MAT-7 national norm sample included private school students and possibly students from classrooms designated as gifted/talented; thus, it differs slightly from the 1999 Texas sample which included public school students from non-gifted/talented classrooms only. Test booklets and answer documents were customized for each grade level and content area tested. While no classroom tested more than one subject area, multiple classrooms may have been selected from some campuses. Form T of the MAT-7 was used to ensure security. This form was developed as a secure form and is printed on demand only. States using Form T must sign an additional security agreement.

Sample Participants

A stratified cluster sample of classrooms not classified as special education nor gifted/talented was selected for each grade tested from Texas public schools to participate in the study. The Texas Education Agency provided Harcourt Educational Measurement (HEM) with summaries of enrollment by grade for the public schools in Texas, including ethnic information. Harcourt Educational Measurement used this information to create the stratified cluster sample of classrooms which met the sample size targets specified in the Introduction.

Details of the stratification and sampling procedure follow. At each grade level, campuses were grouped into one of six sampling regions—(1) Northeast Region, (2) Southeast Region, (3) Central-Valley Region, (4) North Central Region, (5) Northwest Region, and (6) Big 20 (campuses from the twenty largest school districts). The six sampling regions were used as strata, and a random sampling of non-special education and non-gifted/talented classrooms from public schools was performed within each strata for each grade. After the sampling was performed, the number of classrooms selected per campus at each grade was known; at this point, subject areas were randomly assigned. Subject area assignments were spiraled within a grade, so that each subject area was assigned once before a subject area could be assigned a second time. For example, in a campus selected to test three classrooms at grade 5, the first classroom would administer the mathematics test, the second classroom selected would administer the reading test, and the third classroom would administer the mathematics test.

After completion of the sampling process, each selected campus' district was sent materials for 30 students for each selected classroom. Schools were allowed to pick the most convenient testing week between January 11, 1999, and April 30, 1999, and each campus was free to select which of its classrooms would test within a designated grade. The Coordinator Manual instructed the test coordinator to select "typical classrooms." Harcourt Educational Measurement defined "typical classrooms" as classrooms that were not classified as primarily special education or primarily gifted/talented.

Table 1 summarizes the number of districts and campuses that were selected to participate in the study by each grade and content area. It also shows the percent of districts and campuses eligible for participation. The total numbers of districts and campuses are not necessarily the sum of the content area figures since some districts and campuses tested more than one content area.

Number and Percentage of Districts and Campuses Selected to Participate in the 1999 Texas National Comparative Data Study

Grade	Content	Districts		Campuses	
		N	%	N	%
3	Reading	204	19.6	391	11.6
	Math	211	20.2	396	11.8
	Total	253	24.3	573	17.0
4	Reading	229	22.0	410	12.3
	Math	226	21.7	407	12.2
	Language	223	21.4	410	12.3
	Total	333	32.0	856	25.8
5	Reading	206	19.8	361	11.4
	Math	200	19.2	358	11.3
	Total	245	23.6	543	17.2
6	Reading	155	14.9	282	13.0
	Math	158	15.1	285	13.1
	Total	197	18.9	372	17.1
7	Reading	161	15.6	299	16.9
	Math	156	15.1	294	16.6
	Total	199	19.3	397	22.5
8	Reading	183	17.7	308	17.1
	Math	184	17.8	307	17.0
	Language	182	17.6	305	16.9
	Science	190	18.4	310	17.2
	Social Studies	185	17.9	309	17.1
	Total	377	36.5	684	37.9
10	Reading	178	17.8	273	16.8
	Math	173	17.3	271	16.7
	Language	186	18.6	284	17.5
	Total	281	28.1	409	25.1
All	Total	595	55.3	2872	44.4

Table 1

Weighting Procedure and Computation of Means and Standard Errors

In order to ensure that the test results were representative of all students in the state, an ethnic weighting procedure was applied. The Texas Education Agency provided HEM with ethnic enrollment information for each grade at each campus in the state. Based on this information, the proportion of students in each ethnic group considered (African American, Asian American, Hispanic, Native American, and white) was calculated for each sampling region and grade. These figures were used to calculate the ethnic sampling targets for each sampling region (strata), grade, and content area. After the tests were scored, if the actual n-counts failed to closely match the target n-counts, weighting was performed so that the weighted n-counts did closely match the target n-counts. All reported analyses are based on the weighted data. The estimate of each mean NCE was computed as follows:

$$\hat{\bar{Y}} = \left(\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} y_{hij} \right) / w...$$

where H = the number of strata (sampling regions),

n_h = the number of clusters (classrooms) in stratum h,

m_{hi} = the number of measurements (NCE values) in cluster i of stratum h,

y_{hij} = the value of measurement j in cluster i of stratrum h,

w_{hij} = the weight for measurement j in cluster i of stratrum h,

and $w...$ = the sum of all weights over j, i and h.

Taylor series expansion theory was used to estimate the variance of the estimated mean $\hat{\bar{Y}}$. The estimated variance was computed as follows:

$$\hat{V}(\hat{\bar{Y}}) = \sum_{h=1}^{H} \frac{n_h}{n_h - 1} \sum_{i=1}^{n_h} (e_{hi.} - \bar{e}_{h..})^2$$

where $e_{hi.} = \left(\sum_{j=1}^{m_{hi}} w_{hij} (y_{hij} - \hat{\bar{Y}}) \right) / w...$

and $\bar{e}_{h..} = \left(\sum_{i=1}^{n_h} e_{hi.} \right) / n_h$.

The estimated standard error of the mean is the square root of the estimated variance.

Table 2 lists target n-counts, unweighted n-counts, and weighted n-counts.

Texas National Comparative Data Study (1999)
Sample Sizes

	Grade						
	3	4	5	6	7	8	10
Total Reading							
Target n-count	12,508	12,501	12,509	12,507	12,509	12,499	12,500
Unweighted n-count	7,639	7,743	8,609	6,851	6,725	6,792	6,577
Weighted n-count	12,530	12,730	13,042	11,843	12,293	12,299	11,869
Total Mathematics							
Target n-count	12,508	12,501	12,509	12,507	12,509	12,499	12,500
Unweighted n-count	7,650	7,863	8,475	7,129	6,772	6,764	6,789
Weighted n-count	12,815	12,610	13,001	12,033	12,089	12,398	12,172
Language							
Target n-count		12,501				12,499	12,500
Unweighted n-count		7,660				6,827	6,728
Weighted n-count		12,771				12,371	12,023
Science							
Target n-count						12,499	
Unweighted n-count						7,032	
Weighted n-count						12,659	
Social Studies							
Target n-count						12,499	
Unweighted n-count						7,147	
Weighted n-count						12,785	

Table 2

Statewide Results

The following pages of this report provide statewide results of the testing aggregated across all classrooms.

SUMMARY STATEWIDE RESULTS

Table 3 shows the mean national normal curve equivalent (NCE), the associated standard deviation, the sample size, the reliability index (coefficient alpha internal consistency estimate), and the standard error of the mean for each subject area at each grade level.

The NCE is a normalized standard score with a mean of 50 and a standard deviation of 21.06. It is commonly used to present data across years, especially when different test instruments are employed. NCEs have a fixed relationship to percentile ranks, such that a percentile rank of 50 corresponds to an NCE of 50.0, which makes them easy to interpret. It must be noted that the mean NCEs given in Table 3 may be slightly biased estimates of the true mean NCEs for Texas students due to the following sampling differences between the 1997 MAT-7 national norm sample and the 1999 Texas sample: the MAT-7 national norm sample included private school students and possibly students from classrooms designated as gifted/talented; the Texas sample excluded private school students and students from classrooms that were designated as primarily gifted/talented.

The standard error of the mean provides information regarding the variability to be expected in the mean NCE if a different sample of students is obtained. This value is often used to place a confidence interval around the mean.

Figure 1 provides a graphic display of the mean national NCE across all content areas and grades.

Figures 2–6 provide mean national NCE graphs across grades for each content area.

Statewide Results
National NCE Statistics

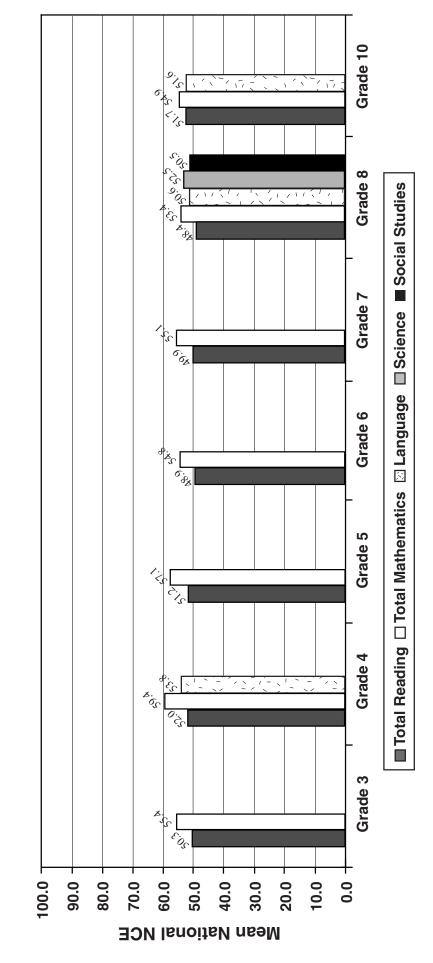|	Grade						
	3	4	5	6	7	8	10
Total Reading							
Mean	50.3	52.0	51.2	48.9	49.9	48.4	51.7
S.D.	18.6	21.2	19.6	24.0	25.1	24.0	24.8
N	12,530	12,730	13,042	11,843	12,293	12,299	11,869
Reliability	0.94	0.95	0.94	0.95	0.94	0.94	0.94
Standard Error of the Mean	0.39	0.45	0.45	0.63	0.65	0.66	0.80
Total Mathematics							
Mean	55.4	59.4	57.1	54.8	55.1	53.4	54.9
S.D.	20.3	19.2	19.9	21.0	22.8	22.5	24.0
N	12,815	12,610	13,001	12,033	12,089	12,398	12,172
Reliability	0.91	0.91	0.92	0.93	0.93	0.93	0.91
Standard Error of the Mean	0.46	0.40	0.46	0.58	0.67	0.64	0.70
Language							
Mean		53.8				50.6	51.6
S.D.		21.0				26.1	24.2
N		12,771				12,371	12,023
Reliability		0.90				0.91	0.90
Standard Error of the Mean		0.41				0.71	0.66
Science							
Mean						52.5	
S.D.						23.8	
N						12,659	
Reliability						0.80	
Standard Error of the Mean						0.52	
Social Studies							
Mean						50.5	
S.D.						21.5	
N						12,785	
Reliability						0.80	
Standard Error of the Mean						0.50	

Table 3

Texas National Comparative Data Study (1999)
Statewide Results

Figure 1

Texas National Comparative Data Study (1999)
Statewide Results

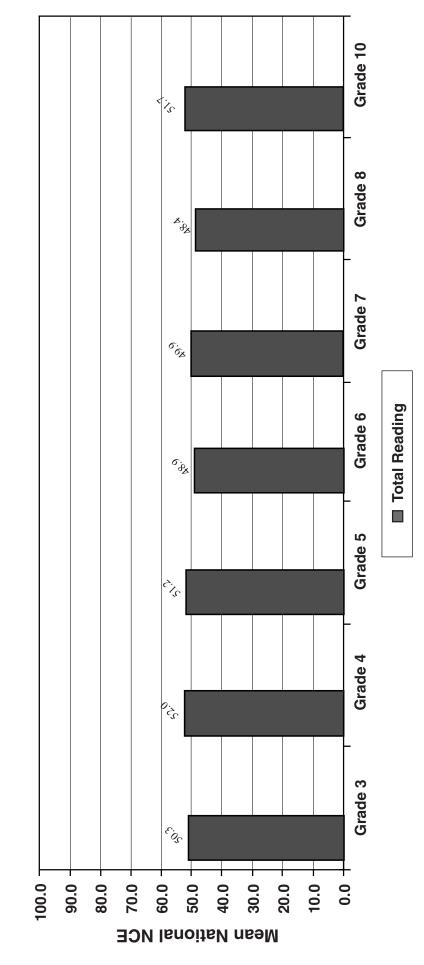**Mean National NCE**

Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 10
50.3	52.0	51.2	48.9	49.9	48.4	51.7

■ Total Reading

Figure 2

**Texas National Comparative Data Study (1999)
Statewide Results**

Mean National NCE

Grade 3: 55.4
Grade 4: 59.4
Grade 5: 57.1
Grade 6: 54.8
Grade 7: 55.1
Grade 8: 53.4
Grade 10: 54.9

☐ Total Mathematics

Figure 3

Texas National Comparative Data Study (1999)
Statewide Results

Mean National NCE

100.0 — 90.0 — 80.0 — 70.0 — 60.0 — 50.0 — 40.0 — 30.0 — 20.0 — 10.0 — 0.0

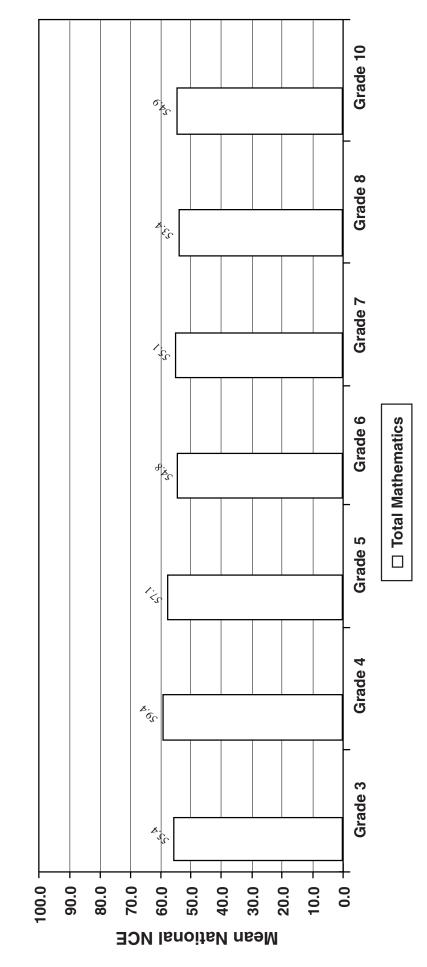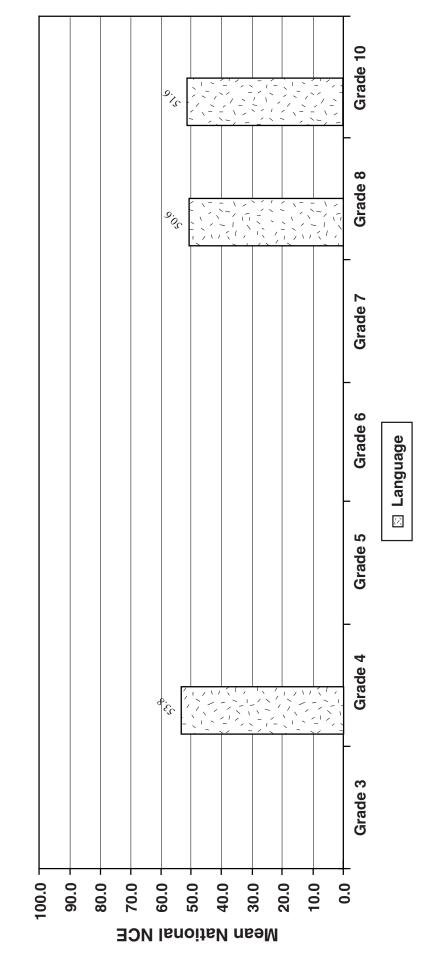| Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 10
	53.8				50.6	51.6

⊡ **Language**

Figure 4

Texas National Comparative Data Study (1999)
Statewide Results

Figure 5

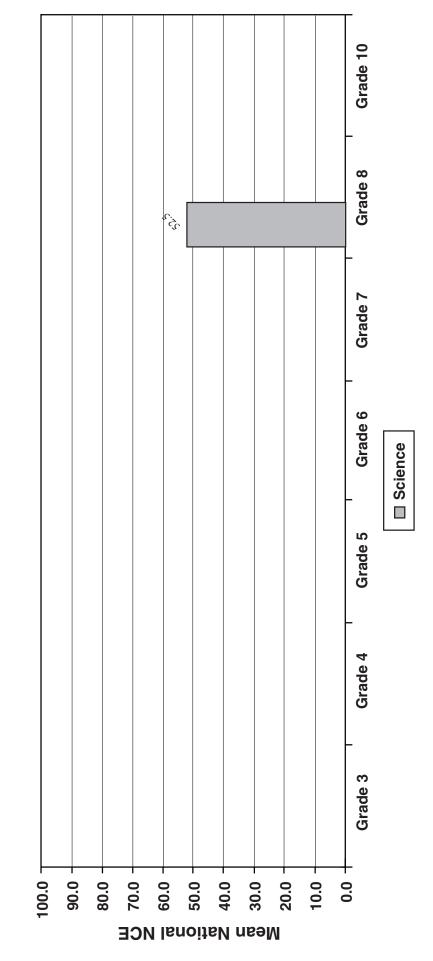**Texas National Comparative Data Study (1999)
Statewide Results**

Mean National NCE

100.0 90.0 80.0 70.0 60.0 50.0 40.0 30.0 20.0 10.0 0.0

Grade 3   Grade 4   Grade 5   Grade 6   Grade 7   Grade 8   Grade 10
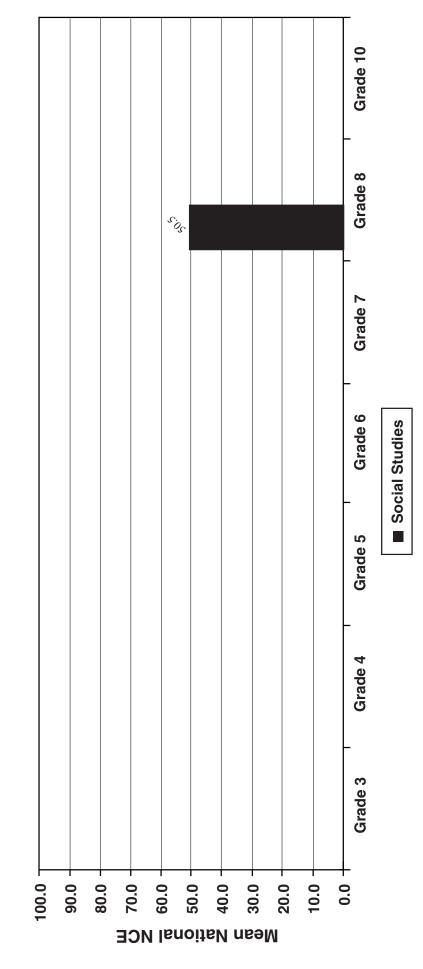
50.5

■ Social Studies

Figure 6

Longitudinal Results

The final section provides a longitudinal analysis incorporating the results of the 1993 statewide Norm-referenced Assessment Program for Texas (NAPT) results, the 1994 voluntary use of the NAPT, the 1995 and 1996 MAT-7 results, and the 1999 MAT-7 results. Please note that due to differences between the NAPT and MAT-7, as well as differences in the grades sampled, not all comparisons can be made. For these analyses, MAT-7 subject areas were paired with NAPT subject areas in the following manner:

MAT-7	NAPT
Total Reading	Reading Comprehension
Total Mathematics	Mathematics Total
Language	Written Expression
Science	Science
Social Studies	Social Studies

Note that there is not an associated mathematics total score for NAPT at grade 10 since the NAPT at this grade was not configured to report a total mathematics score.

Table 4 shows mean National NCEs for each year within each content area. Figures 7 through 11 provide a graphic representation of the mean national NCE for each content area at each grade. Comparisons of changes in specific values across the years should be made with caution, however, for the following reasons: (1) the national norm sample for the 1995 and 1996 MAT-7 was different than the national norm sample for the 1999 MAT-7, and both national norm samples were different than the national norm sample for the 1993 and 1994 NAPT, and (2) the differences in the test items comprising each content area. Thus, longitudinal comparisons should only be used to identify general trends.

Texas National Comparative Data Study (1999)
Longitudinal Results
Mean National NCE

	Grade						
	3	4	5	6	7	8	10
MAT-7: Total Reading **NAPT: Reading Comprehension**							
MAT-7 1999	50.3	52.0	51.2	48.9	49.9	48.4	51.7
MAT-7 1996	49.9	49.3	50.2	48.2	47.7	47.2	44.5
MAT-7 1995	48.9	47.2	48.4	47.3	48.4	46.3	43.1
NAPT 1994	50.6	51.8	49.7	50.2	49.9	49.9	47.7
NAPT 1993	50.4	51.8	49.7	49.7	49.7	48.4	48.4
MAT-7: Total Mathematics **NAPT: Mathematics Total**							
MAT-7 1999	55.4	59.4	57.1	54.8	55.1	53.4	54.9
MAT-7 1996	54.7	56.5	57.1	54.2	53.3	50.1	47.4
MAT-7 1995	58.0	58.9	58.2	53.3	51.9	52.1	51.3
NAPT 1994	55.9	54.7	52.6	50.9	49.3	50.2	
NAPT 1993	55.0	54.4	51.0	50.0	49.1	49.2	
MAT-7 Language **NAPT: Written Expression**							
MAT-7 1999		53.8				50.6	51.6
MAT-7 1996		55.8				54.1	48.9
MAT-7 1995		54.9				52.0	47.9
NAPT 1994	49.7	54.4	52.3	53.4	53.5	52.3	50.9
NAPT 1993	50.6	54.6	51.7	52.9	53.2	51.5	51.9
MAT-7 Science **NAPT: Science**							
MAT-7 1999						52.5	
MAT-7 1996						50.7	
MAT-7 1995						52.8	
NAPT 1994	52.4	51.3	52.9	52.0	51.5	51.3	52.8
NAPT 1993	50.9	51.4	51.5	50.4	49.3	48.5	50.0
MAT-7 Social Studies **NAPT: Social Studies**							
MAT-7 1999						50.5	
MAT-7 1996						51.7	
MAT-7 1995						50.9	
NAPT 1994	53.7	50.8	49.4	52.3	51.4	51.9	53.2
NAPT 1993	52.1	51.1	49.4	50.7	49.3	49.4	50.2

Table 4

Texas National Comparative Data Study (1999)
Longitudinal Results
Total Reading

Mean National NCE

Grade 3 — 50.4, 50.6, 48.9, 49.9, 50.3
Grade 4 — 51.8, 47.2, 49.3, 52.0
Grade 5 — 49.7, 48.4, 50.2, 51.2
Grade 6 — 49.7, 50.2, 47.3, 48.2, 48.9
Grade 7 — 49.7, 48.4, 47.7, 49.9
Grade 8 — 48.4, 46.3, 47.3, 48.4
Grade 10 — 48.4, 47.7, 43.1, 44.5, 51.7

NAPT 1993 ☐ NAPT 1994 NAPT 1994 ☐ NAPT 1994 MAT-7 1995 ☐ MAT-7 1995 MAT-7 1996 ■ MAT-7 1999

Figure 7

Texas National Comparative Data Study (1999)
Longitudinal Results
Total Mathematics

Mean National NCE

Grade 3 Grade 4 Grade 5 Grade 6 Grade 7 Grade 8 Grade 10

NAPT 1993 ☐ NAPT 1994 ☒ MAT-7 1995 ☐ MAT-7 1996 ■ MAT-7 1999

Figure 8

Texas National Comparative Data Study (1999)
Longitudinal Results
Language

Mean National NCE

Grade 3 — 50.6, 49.7
Grade 4 — 54.6, 54.4, 54.9, 55.8, 53.8
Grade 5 — 51.7, 52.3
Grade 6 — 52.9, 53.4
Grade 7 — 53.2, 53.5
Grade 8 — 51.5, 52.3, 52.0, 54.1, 50.6
Grade 10 — 51.9, 50.9, 47.9, 48.9, 51.6

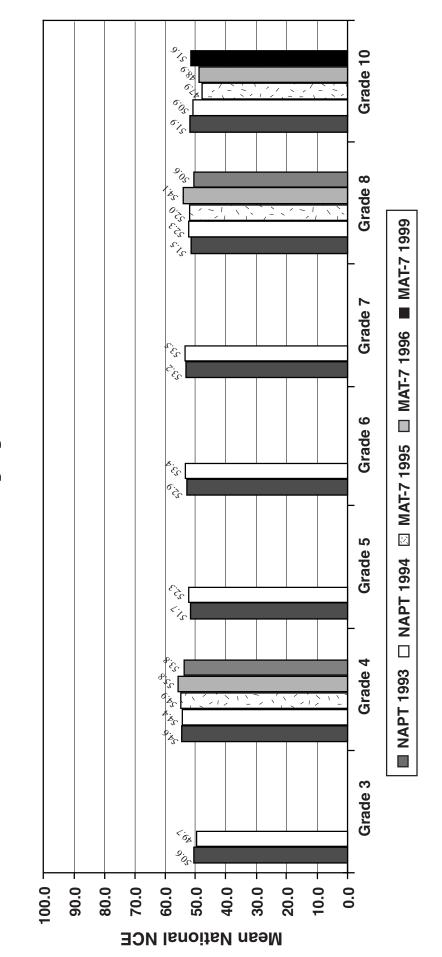Legend: NAPT 1993 ■ NAPT 1994 □ MAT-7 1995 ▨ MAT-7 1996 ■ MAT-7 1999 ■

Figure 9

Texas National Comparative Data Study (1999)
Longitudinal Results
Science

Figure 10

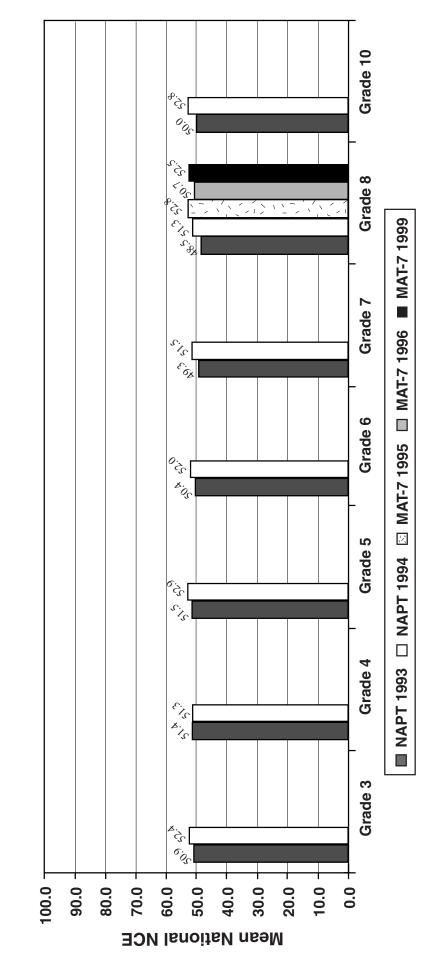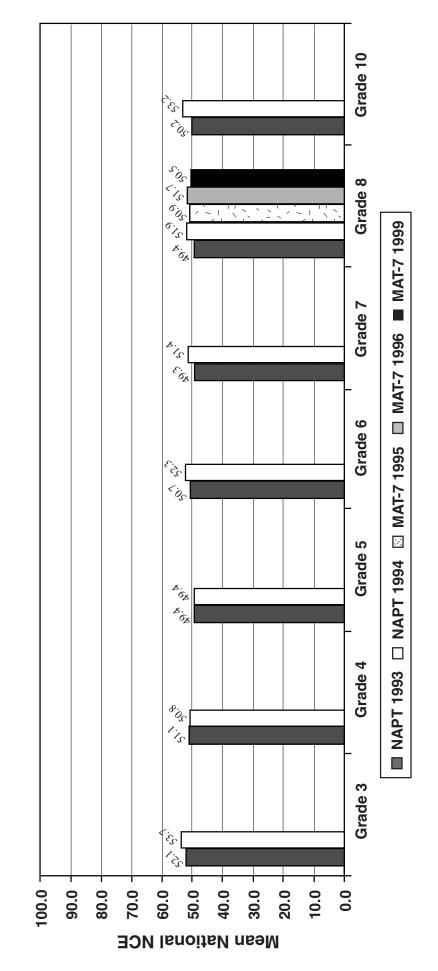**Texas National Comparative Data Study (1999)**
Longitudinal Results
Social Studies

Mean National NCE

Legend: NAPT 1993 ■ NAPT 1994 □ MAT-7 1995 ▨ MAT-7 1996 ▨ MAT-7 1999 ■

Grade 3: 52.1, 53.7
Grade 4: 50.8, 51.1
Grade 5: 49.4, 49.4
Grade 6: 50.7, 52.3
Grade 7: 49.3, 51.4
Grade 8: 49.4, 51.9, 50.9, 51.7, 50.5
Grade 10: 50.2, 53.2

Figure 11