

Background and Interpretation of 2005 National Assessment of Educational Progress (NAEP) Results for Maine in Mathematics and Reading

J. H. Kennedy, NAEP State Coordinator for Maine

Contents of this Report

- National Results: The Differences Between Long Term Trend NAEP and Main (NCLB) NAEP
- Maine Educational Assessment (MEA) vs. NAEP: Comparing Achievement, Balance of Representation, and Trend Lines
- No Child Left Behind in Maine: Disaggregation of Maine NAEP Data
- Appendix A: Comparisons of NAEP and MEA Achievement Level Results

Note: A summary of significant 2005 NAEP results is given in the 2005 NAEP Press Release.

National Results

- NCLB NAEP vs. Long Term Trend NAEP

Earlier this year, NAEP released results of the 2004 Long Term Trend Study in Reading and Mathematics. Below is an analysis of the differences between that assessment and the assessment for which results are reported here. The 2005 NAEP, often referred to as the 'main' MAEP, is the assessment for which state participation is mandated under the *No Child Left Behind* Act. The following is a screen print from the Internet, and the links are not active.

What Are the Differences Between Long-Term Trend NAEP and Main NAEP?

Although long-term trend and main NAEP both assess mathematics and reading, there are four main differences—the content assessed, the students selected (sampled), the administration timing, and the results reported. These differences mean that results from long-term trend (LTT) and main NAEP cannot be compared directly, although comparisons of the patterns over time of the two assessments, especially for student demographic groups, may be informative, keeping in mind the content differences.

Learn more about the [long-term trend assessment](#).

Learn more about the main NAEP [national and state](#) assessments.

	Long-Term Trend Assessment	Main NAEP Assessment
Purpose	Measures student performance in mathematics and reading every four years. Last reported for 2004; will be reported next for 2008.	Measures student performance in mathematics and reading every two years, most recently in 2005. Other subjects are also assessed.

Content Assessed	Has remained essentially unchanged since first administration	Changes about every decade to reflect changes in curriculum in the
-------------------------	---------------------------------------------------------------	--------------------------------------------------------------------

	<p>(1971 for reading, 1973 for mathematics), although some changes were initiated in 2004. Note: Questions and assessment instrument are very different from main NAEP; the plan for LTT assessment has been static, whereas frameworks for main NAEP change.</p> <p>Reading features shorter passages, and focuses on locating specific information, making inferences, and identifying the main idea of a passage. Students respond to questions in multiple-choice format; there are also a few questions requiring an extended answer (usually scored on a five-point scale).</p> <p>Mathematics focuses on basic computational skills in four content areas: numbers and operations, measurement, geometry, and algebra. Students respond to questions in multiple-choice format; there are also a few short answer (scored on a two- or three-point scale) and a few extended answer (usually scored on a four-point scale). Students are not asked to show or explain their work.</p>	<p>nation's schools. New frameworks reflect these changes.</p> <p>Reading requires students to read longer passages or pairs of passages; measures a range of reading skills, from identifying explicitly stated information, to making complex inferences about themes, to comparing multiple texts on a variety of dimensions. Students respond to questions of three possible types: multiple choice, short answer (scored on a two- or three-point scale), and extended answer (scored on a four-point scale).</p> <p>Mathematics focuses on five content areas: number properties and operations, measurement, geometry, data analysis and probability, and algebra. Students respond to questions of several possible types: multiple choice, short answer, and extended answer. There are three answer categories for constructed-response questions (right or wrong; right, partially right, or wrong; or extended constructed-response with several levels of partial credit, scored on a four-point scale). Students may be asked to explain their work.</p>
--	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Students Sampled	<p>Selected by age (9, 13, and 17) to represent nation. Smaller sample sizes than in main NAEP restrict comparisons to main demographic groups. In 2004, results could be reported only for White, Black, and Hispanic students, and for public schools.</p>	<p>Selected by grade (4, 8, and 12). Students represent nation in even-numbered years, but represent participating states and urban districts in odd-numbered years. Larger sample sizes usually permit reporting results for smaller minorities, such as Asians and Pacific Islanders, and American Indians.</p>
Administration	<p>Assessment every four years, throughout the school year: in October through December for 13-year-olds, January through March for 9-year-olds, and March through May for 17-year-olds. See schedule for all assessments.</p>	<p>Mathematics and reading every two years, but assessment of other subjects requires administration of main NAEP each year from late January through early March. Main NAEP administered only at national level in even-numbered years; participating states and selected urban jurisdictions are represented in odd-numbered years.</p>

Results Reported	Provides results on performance and how it has changed over time. Reported as performance levels using scale scores. Does not provide results in terms of achievement levels. Report provides only national results.	Provides results on performance and how it has changed over time using scale scores, but also reports achievement levels (<i>Basic, Proficient, and Advanced</i>). Results have been produced for the nation and the participating states and other jurisdictions since 1990, and for selected urban districts (on a trial basis) since 2002.
-------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Source: NAEP Network

- The 2005 NAEP was largest National sample ever, creating finer distinctions than in the past: 338,000 4th graders and 321,000 8th graders were sampled. Because an increased sample size leads to a decreased margin of errors, small differences in scores may now be statistically significant. See **Statistical Significance** below.
- At 4th grade, Mathematics and Reading scores for the Nation as a whole were up from 2003. At 8th grade, Mathematics scores were up and Reading scores were down for the Nation as a whole.

Maine Educational Assessment (MEA) vs. NAEP

Several national organizations, including Achieve.org and Schoolresults.org have used various methods to compare state and National assessment results. Maine has typically been shown to have a state assessment of comparable rigor to the National Assessment. Below is ranking of the six states each in Reading and Mathematics that had the smallest gap between percentages of students meeting state standards on the state assessment and achieving *At or Above Proficient* on NAEP. For an explanation of the NAEP Achievement Levels, please see NAEP Newsletter #1 (*What is NAEP?*).

Difference Between Students Meeting State Standards & At or Above Proficient on NAEP

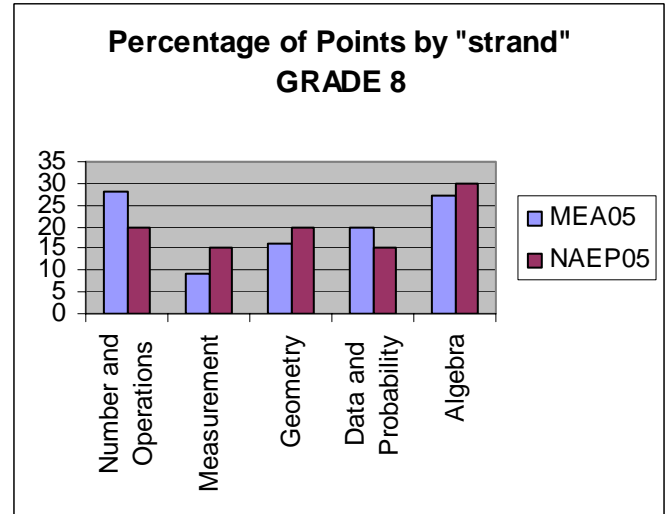
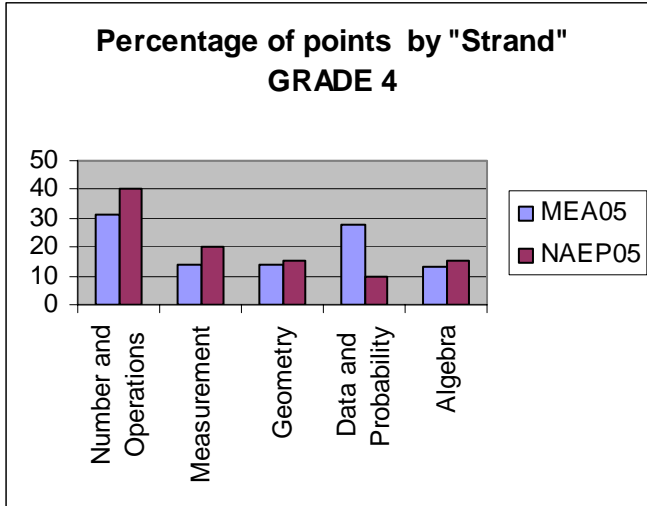
READ			MATH		
RANK	STATE	GAP	RANK	STATE	GAP
1	MS	0	1	ME	-6
2	SC	5	2	VT	-2
3	WY	10	3	WY	-2
4	ME	13	4	MA	-1
5	MA	16	5	HI	1
6	CA	18	6	SC	2

Education Week: "State of the States," 2005

Please see appendix A for comparisons of NAEP and MEA achievement level results.

Below is an analysis of the Balance of Representation in MEA Mathematics and NAEP Mathematics done in the Maine Department of Education. It is important to note that the MEA measures grade spans, while NAEP is designed to assess the achievement of students at particular points in time. Measurement is represented less in the MEA than in NAEP because the Mathematics specialists at the Maine Department of Education believe that this strand should include more hands on activities and use of tools, which is best assessed in the Local Assessment System (LAS). On the other hand, Data and

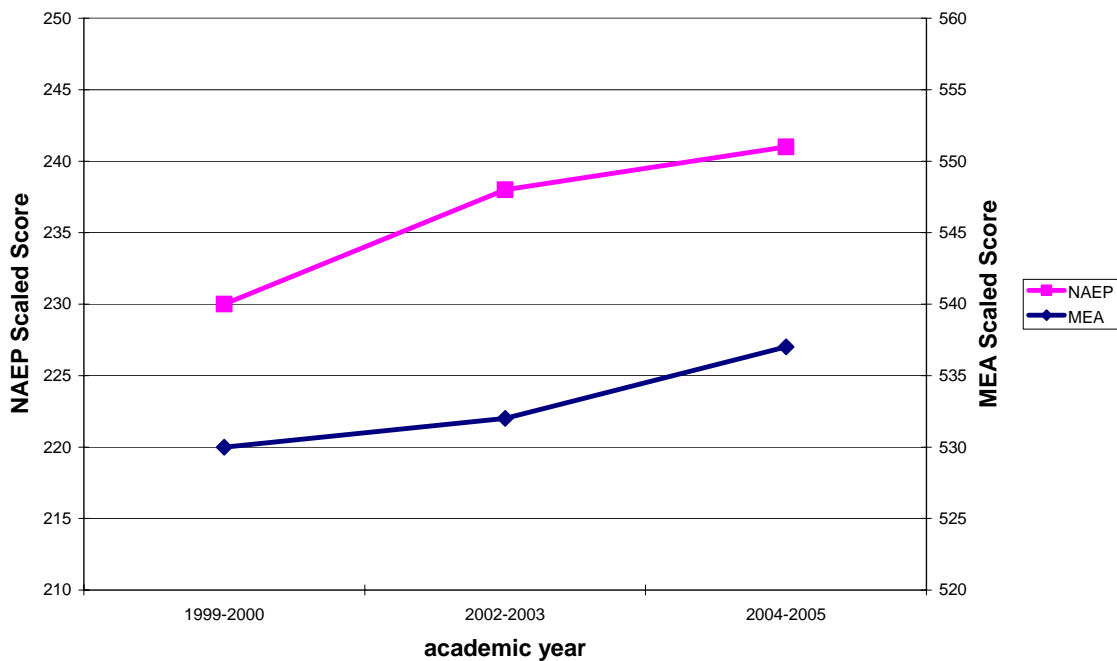
Probability are represented more in the MEA than in NAEP because the Maine Learning Results reflect higher expectations for 4th graders in the area of data and it those areas receive more test emphasis based on Maine's Balance of Representation for the Learning Results.



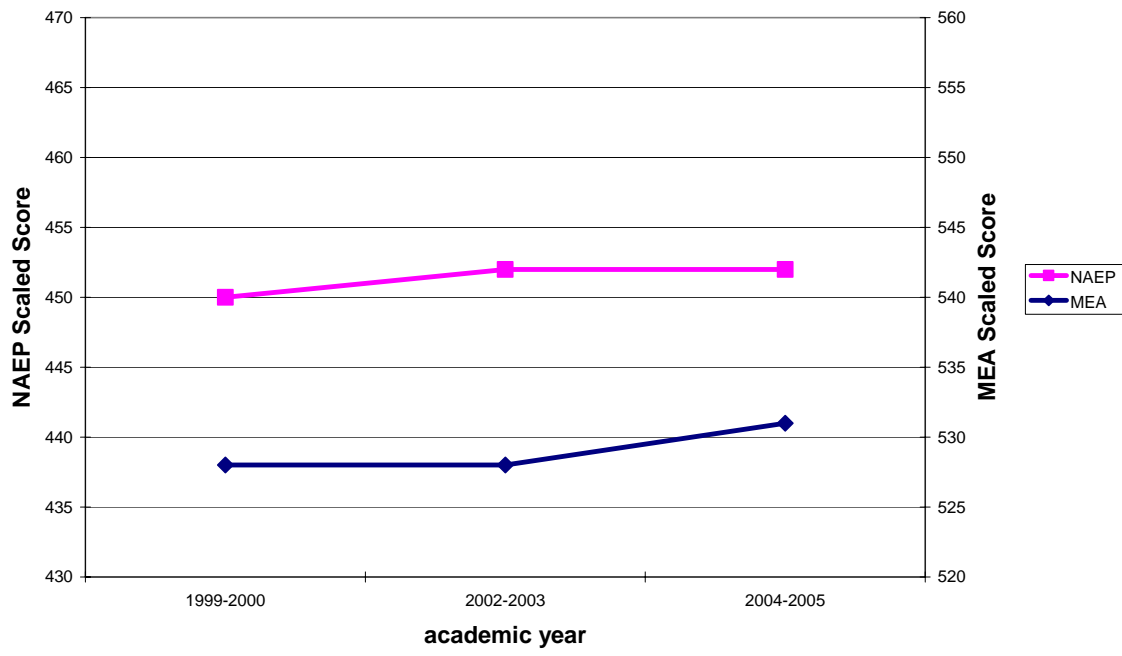
Tad Johnston, MEDOE

The following charts compare performance trends on the MEA and NAEP:

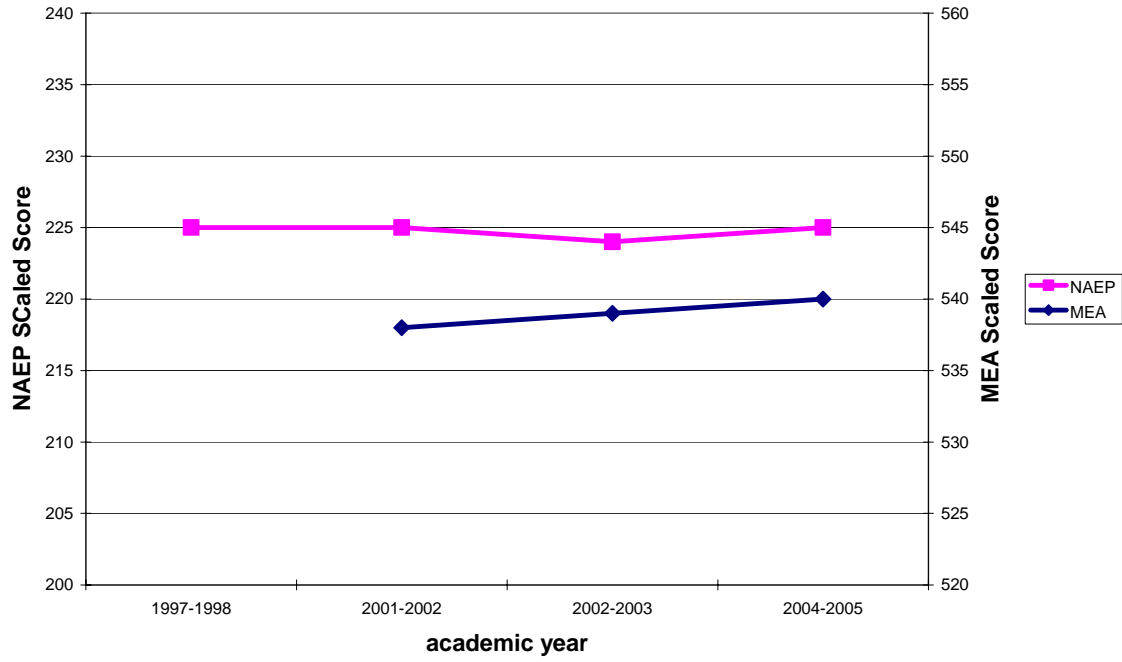
Maine NAEP and MEA 5-Year Trend Grade 4 Mathematics



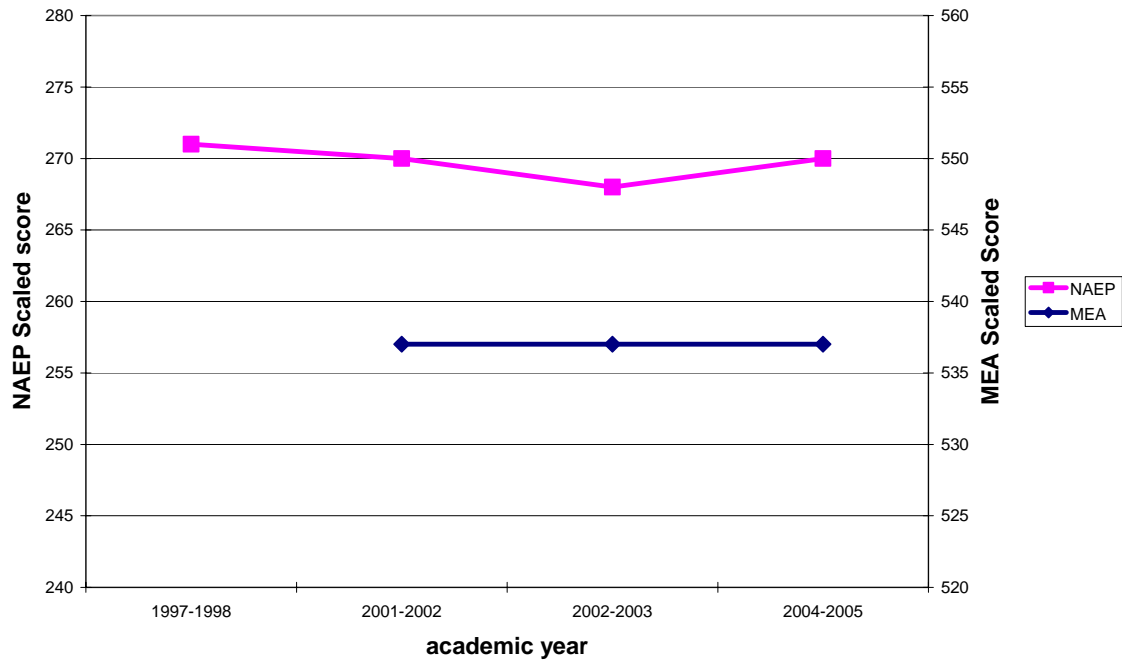
Maine NAEP and MEA 5-Year Trend Grade 8 Mathematics



MAINE NAEP 7-Year Trend with MEA Scores Grade 4 Reading



Maine NAEP 7-Year Trend with MEA Scores Grade 8 Reading



No Child Left Behind in Maine

- Grade 4 – Mathematics scores up; Reading scores stable compared to 2003
- Grade 8 – Mathematics and Reading scores stable compared to 2003

Digging below the surface of these data reveals comparisons state-state and state-Nation comparisons favorable to Maine, with two exceptions:

- Rural 4th grade students are showing more sustained improvement in Reading than in Mathematics.
- Male students at 4th and 8th grade are suppressing Maine Reading scores.

Comparing Whites across states is misleading because of wide range in socioeconomic status for this group. Maine's NAEP sample was 97% White in 2005, while the Nation's sample was 57%. Comparisons of students performance in the lower 10 percent of their cohort, school location, eligibility for free or reduced price lunch, and students with disabilities are more meaningful and are generally quite favorable for Maine.

Suburban subgroup comparisons favor Maine, but that subgroup is relatively small here. Suburban scores are typically the highest for school location in NAEP. In New Hampshire, the larger percentage of suburban students sampled probably contributed to their higher score in 4th Grade Reading.

Maine has the second highest percentage of rural students in the nation. The difference between rural and urban subgroups within Maine was not significant for either subject at either grade. Bangor, Lewiston, and Portland –all having culturally diverse student populations – were sampled extensively in the 2005 NAEP.

The data suggest that in Maine, students tend to enjoy the same educational opportunities regardless of socioeconomic status, school location, or ethnic grouping. See the upcoming *NAEP Newsletter #9* for details of visits to Maine's urban, island, rural and Indian schools during and after the 2005 NAEP administration.

Mathematics

- SPED scores for Mathematics at 4th grade showed significant increases in 2005 and remain higher than for the Nation as a whole:

SPED	4 th	Maine	Nation	<i>Mathematics</i>
2000		210	198	
2003		215	214	
2005		222	218	

- With increased inclusion with the introduction of accommodations in 1998, 4th grade scores began steady increase at all percentiles reported (75th, 50th, 25th, and 10th), indicating that inclusion has served to enhance rather than suppress overall educational progress for Maine.

New Source for These Data: The NAEP Data Explorer

With the release of NAEP 2005 Reading and Mathematics results for the Nation and the states, the National Assessment Governing Board (NAGB) also introduced a new tool for investigating the results, The NAEP Data Explorer (NDE). The NDE replaces the NAEP Data Tool which provided only pre-calculated data

based upon single variables, while the NDE is capable of calculating results for multiple variables (rural males participating in the free and reduced school lunch program, for instance) in real time. All Maine results presented here were produced at the Maine Department of Education using the NDE and subjected to a test of statistical significance provided by the NDE.

Since NAEP results are predictions of student performance for identified groups based upon a sampling of only a portion of those groups, the average scaled scores and percentages of students at achievement levels are subject to a margin of error. If a difference between specific scores or percentages is calculated by the NDE to be not significant, this means that there is no predictable difference in the performance of the two groups whose results are being examined or of the single group over time.


A screen print of the significance test output from the NAEP Data Explorer for 2000-2005 NAEP SPED results in Maine is reproduced below. Three thumbnails prints of significance tests for SPED results in 8th grade Mathematics and Reading at both grades follow, indicating no demonstrable change in performance; however, Maine SPED scores in these subjects and at these grades remain above the National average for the same subgroup.

The NAEP Data Explorer (NDE) is available to the public at this website. Scroll to the bottom of the Maine NAEP homepage, and click on "NAEP e-Research Center". The NAEP State Coordinator is available to conduct on-site workshops in the use of the NDE. Contact J. H. Kennedy at 207-624-6636 (john.kennedy@maine.gov) for more information.

Check Statistical Significance

[Print/Save/Export](#)

[Get help](#)

Make selections in the left-hand area below. You will see a preview of your statistical table on the right. The  symbol points to selections that you must make in order to perform a statistical test. When you have made all the necessary selections this symbol disappears, and you can click on the "Compute" button.

Selections

Jurisdiction

Year

Select All

2000

2003

2005

Variable

Student

classified as having a disability

Select All

Yes

No

Statistic Type


Display Options

Significance Test Results

To see how one value compares with the others, read across the row for that value. The displayed symbols indicate whether that value is significantly higher, significantly lower, or not significantly different than the value associated with that column. In some case the significance test may have not been possible for statistical reasons.

NAEP Mathematics Grade 4 - Mathematics
Difference in Average Scale Score Between Years
for selected values
for Student classified as having a disability [IEP] = Yes
Maine

	2000	2003	2005
2000		=	<
2003	=		<
2005	>	>	

 No test was performed

 Significantly lower

 Significantly higher

= No significant difference

¹ Accommodations were not permitted for this assessment

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000, 2003 and 2005 Mathematics Assessments.

[close window](#)

Significance Test Results

To see how one value compares with the others, read across the row for that value. The displayed symbols indicate whether that value is significantly higher, significantly lower, or not significantly different than the value associated with that column. In some cases the significance test may have not been possible for statistical reasons.

NAEP Mathematics Grade 8 Mathematics
Difference in Average Scale Score Between Years
for selected values
for Student classified as having a disability [IEP] = Yes
Maine

	2000	2003	2005
2000		=	=
2003	=		=
2005	=	=	

No test was performed
< Significantly lower
> Significantly higher
= No significant difference

¹ Accommodations were not permitted for this assessment

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000, 2003 and 2005 Mathematics Assessments.

Significance Test Results

To see how one value compares with the others, read across the row for that value. The displayed symbols indicate whether that value is significantly higher, significantly lower, or not significantly different than the value associated with that column. In some cases the significance test may have not been possible for statistical reasons.

NAEP Mathematics Grade 8 Mathematics
Difference in Average Scale Score Between Years
for selected values
for Student classified as having a disability [IEP] = Yes
Maine

	2000	2003	2005
2000		=	=
2003	=		=
2005	=	=	

No test was performed
< Significantly lower
> Significantly higher
= No significant difference

¹ Accommodations were not permitted for this assessment

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000, 2003 and 2005 Mathematics Assessments.

Significance Test Results

To see how one value compares with the others, read across the row for that value. The displayed symbols indicate whether that value is significantly higher, significantly lower, or not significantly different than the value associated with that column. In some cases the significance test may have not been possible for statistical reasons.

NAEP Reading Grade 8 - Reading
Difference in Average Scale Score Between Years
for selected values
for Student classified as having a disability [IEP] = Yes
Maine

	1998	2002	2003	2005
1998		=	=	=
2002	=		=	=
2003	=	=		=
2005	=	=	=	

No test was performed
< Significantly lower
> Significantly higher
= No significant difference

¹ Accommodations were not permitted for this assessment

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1998, 2002, 2003 and 2005 Reading Assessments.

NOTE: The NAEP Mathematics scale ranges from 0 to 500. Observed differences are not necessarily statistically significant.

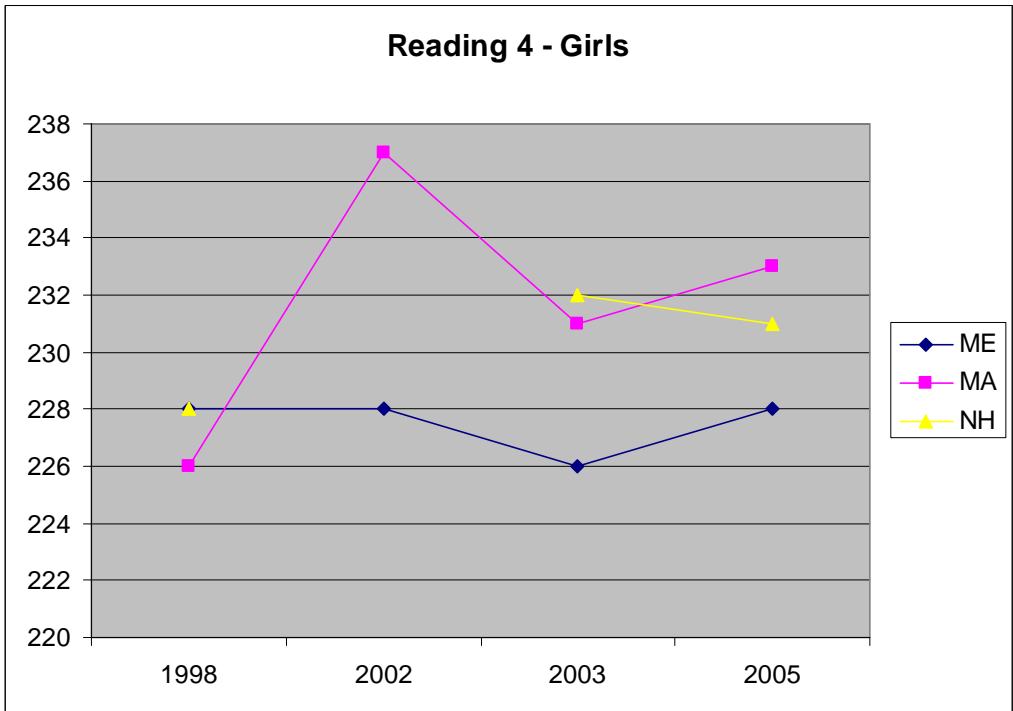
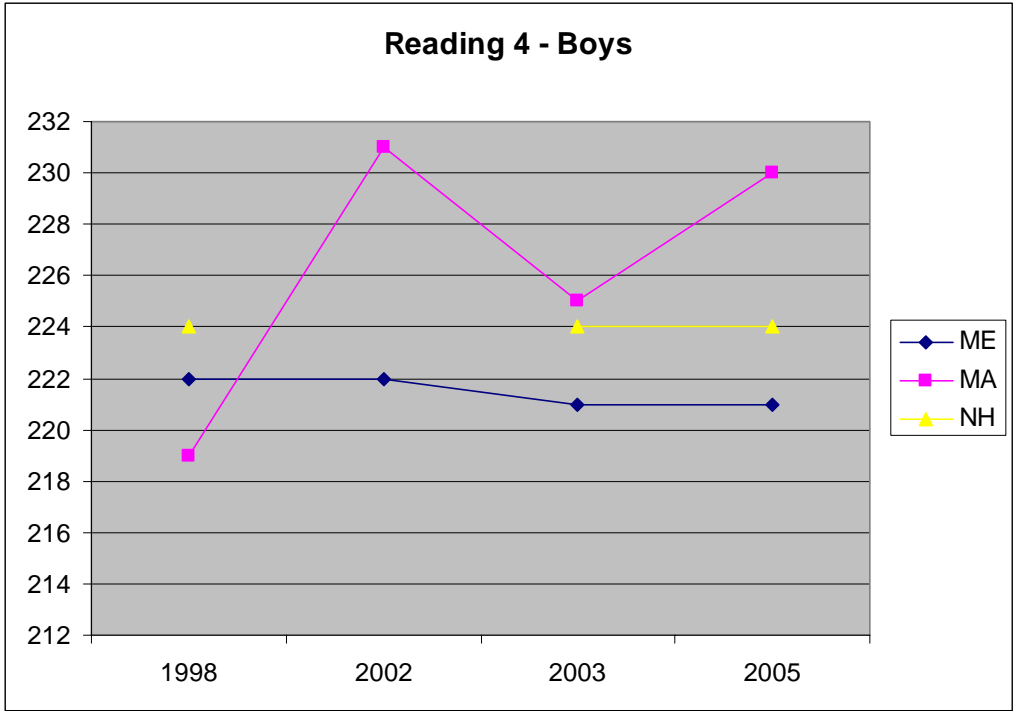
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1992, 1996, 2000, 2003 and 2005 Mathematics Assessments.

The Mathematics scores of the 10th percentile (the lowest achieving students) were significantly higher in Maine than in the Nation as a whole. At 4th grade, Maine students in the 10th percentile had an average score of 206, compared to 199 for students in the 10th percentile nationwide. At 8th grade, Maine students in the 10th percentile had an average score of 238, compared to 230 for students in the 10th percentile nationwide.

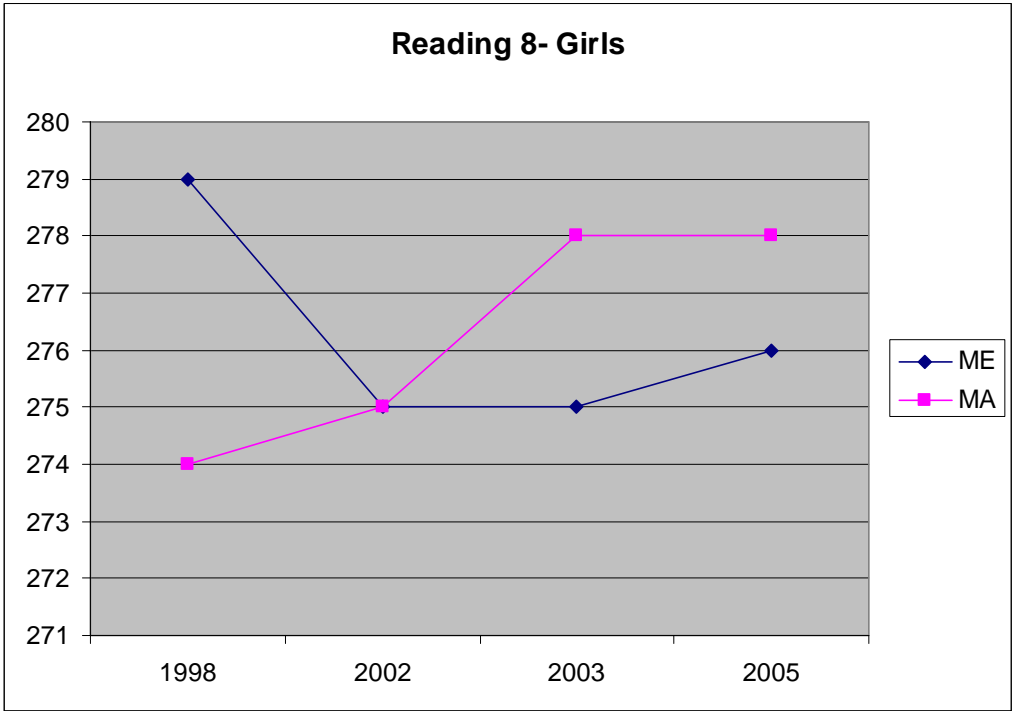
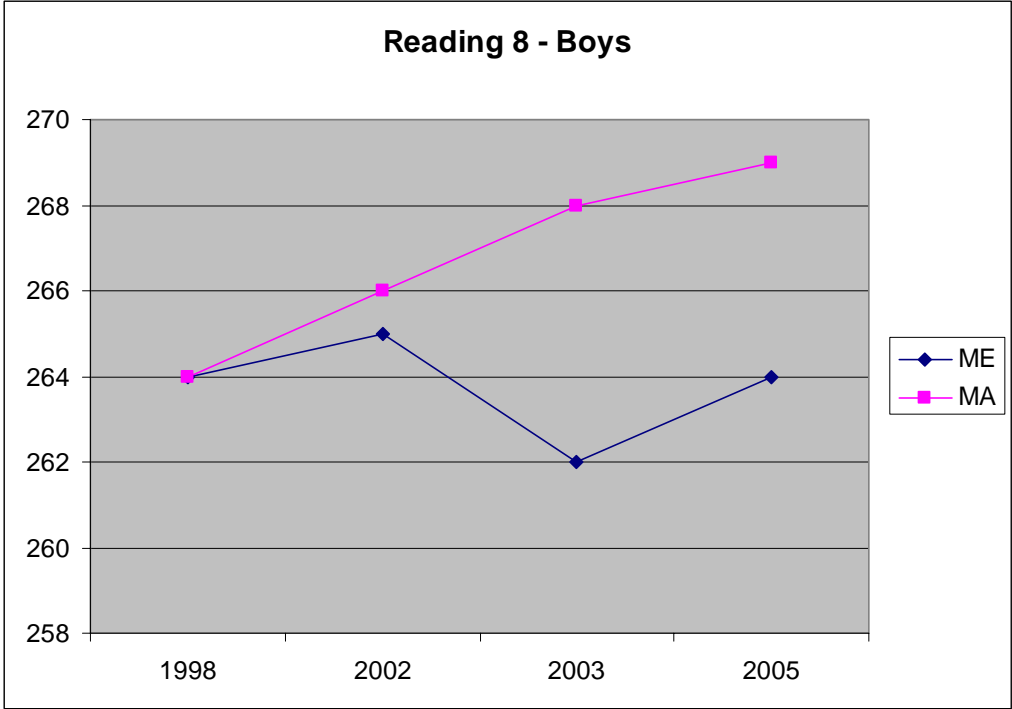
Reading

- Maine's reading results have been stable for 4th grade and have declined slightly for 8th grade since 1998 but have been stable in this century.
- Improvement of male reading scores has been a significant source of Massachusetts' steady increase in scores. Geographic distributions appear to favor New Hampshire, but a strong early childhood literacy effort on the part of human services in the state also appears to be a factor also. Similarly, Massachusetts has undertaken a number of costly initiatives in Reading.
- The line charts below compare trends for 4th grade males and females since 1998.
- Comparison of students eligible for free or reduced price lunch across the nation shows no states with higher scores at 4th grade and most all but five states with lower scores at 8th grade.
- Comparisons of Achievement Level trends for males and females in Maine and Massachusetts show the percentage of Grade 4 girls in Maine *At or Above Proficient* remaining stable and that of the boys declining slightly. At Grade 8, the boys in Maine have remained more or less stable and the girls have recovered slowly from a large decrease in percentage *At or Above Proficient* between 1998 and 2002.

By contrast, the Massachusetts percentages of 4th grade students of both gender peaked in 2002, declined slightly in 2003, and rose in 2005. Percentages for Massachusetts Grade 8 students of both genders have risen consistently since 1998.

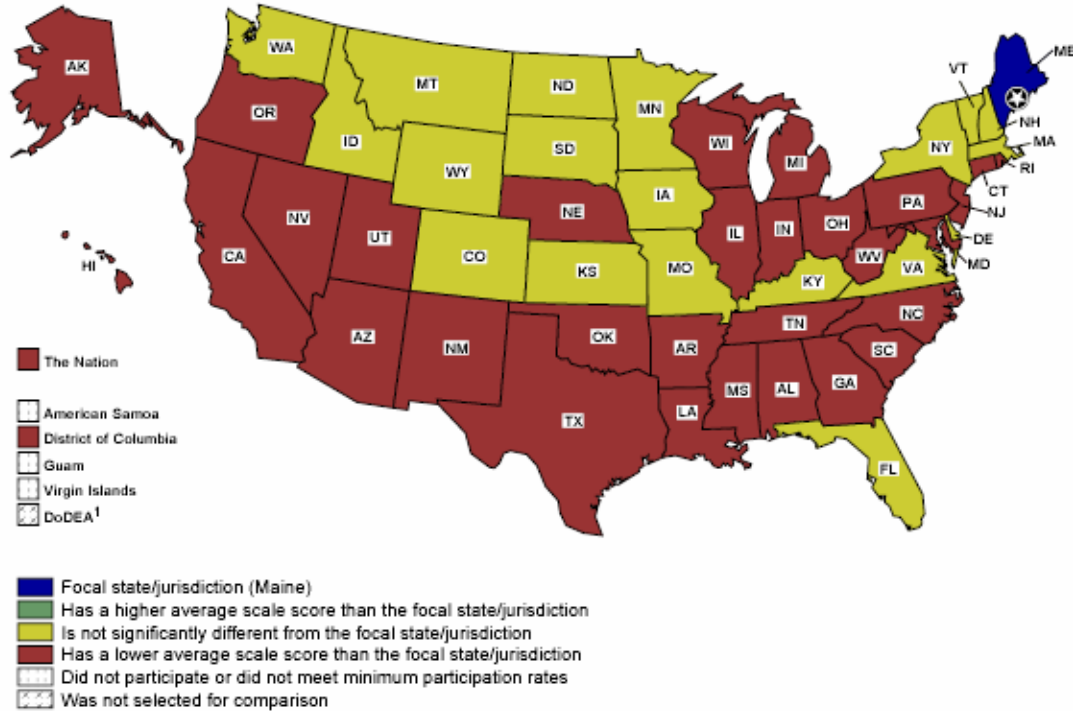


Note: Both NH and ME had 228 in 1998; NH did not participate in 2002



NAEP Reading Grade 4 - Reading
 Difference in Average Scale Score Between Jurisdictions
 For Natl School Lunch Prog eligibility (3 categories) [SLUNCH3] = Eligible
 2005

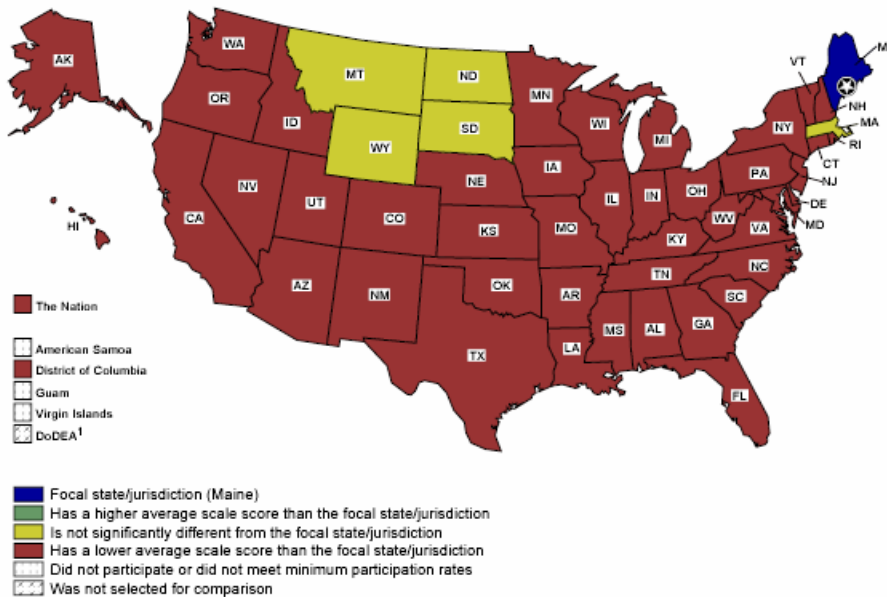
Color



¹ Department of Defense Education Activity schools (domestic and overseas).
 SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Reading Assessments.

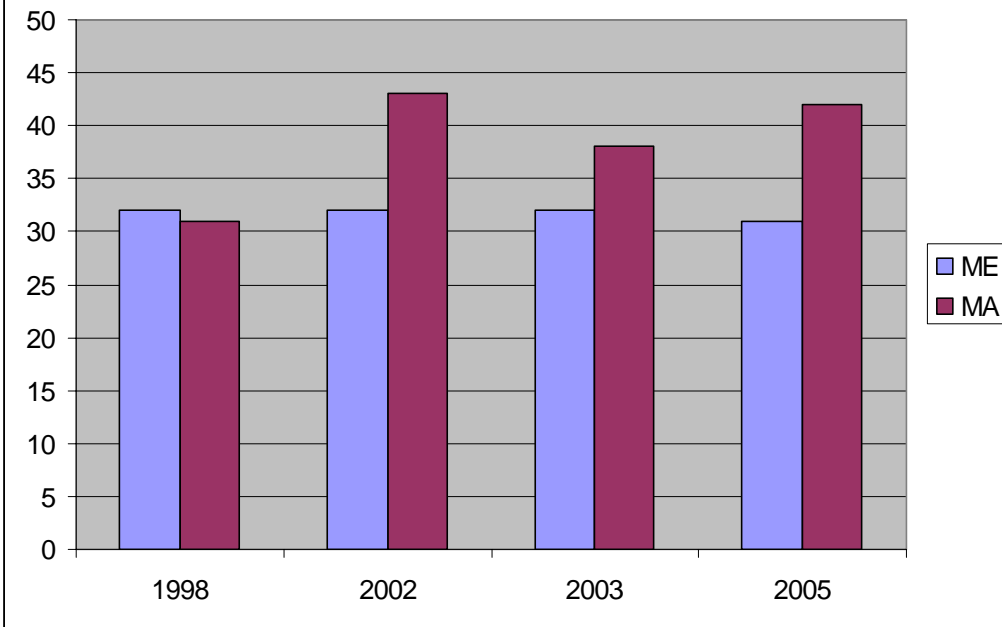
NAEP Reading Grade 8 - Reading
 Difference in Average Scale Score Between Jurisdictions
 For Natl School Lunch Prog eligibility (3 categories) [SLUNCH3] = Eligible
 2005

Color

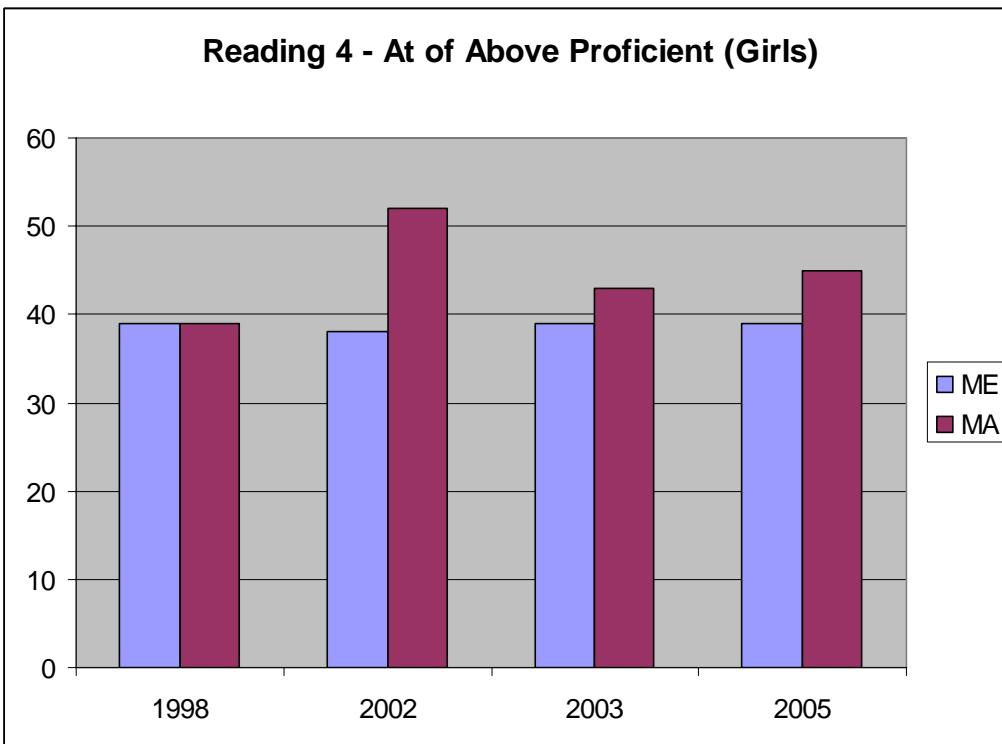


¹ Department of Defense Education Activity schools (domestic and overseas).
 SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2005 Reading Assessments.

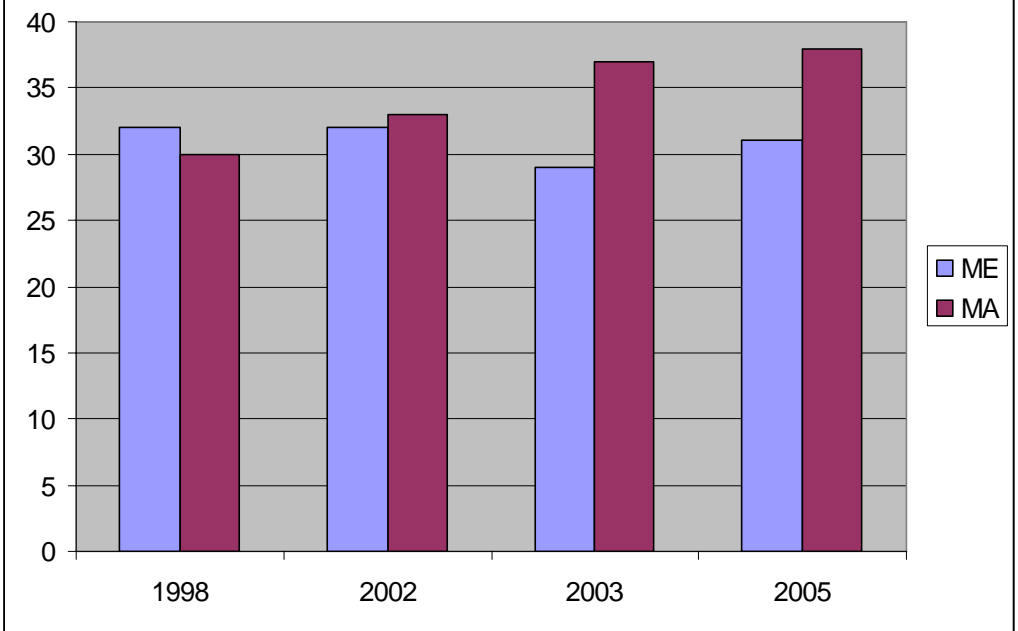
Reading 4 - At or Above Proficient (Boys)



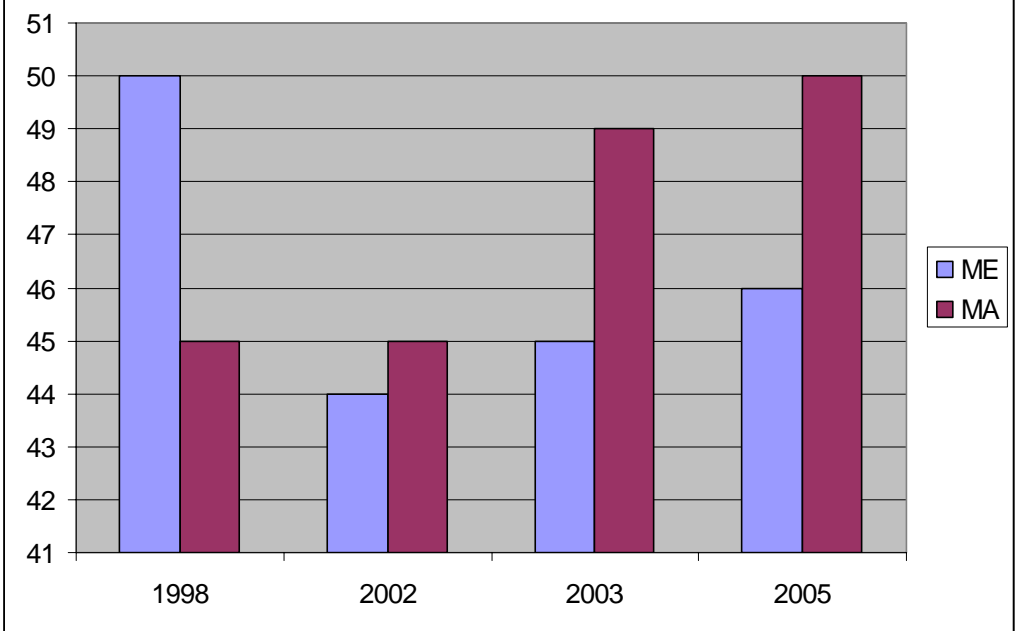
Reading 4 - At of Above Proficient (Girls)



Reading 8 - At or Above Proficient (Boys)



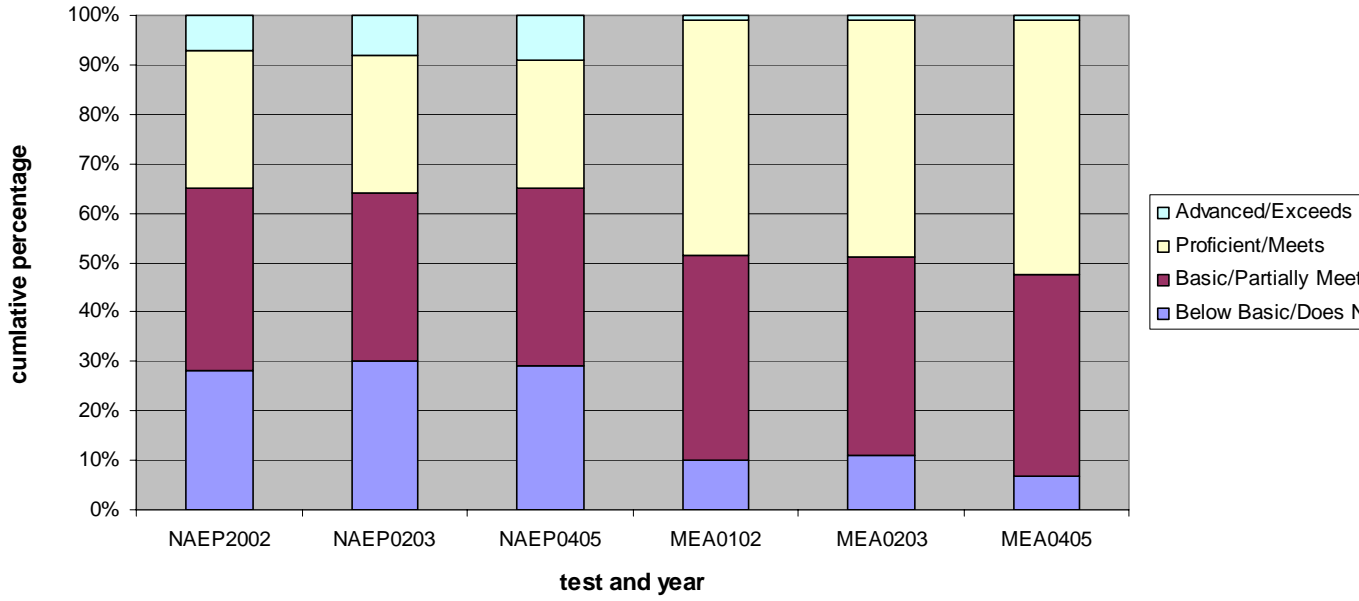
Reading 8 - At or Above Proficient (Girls)



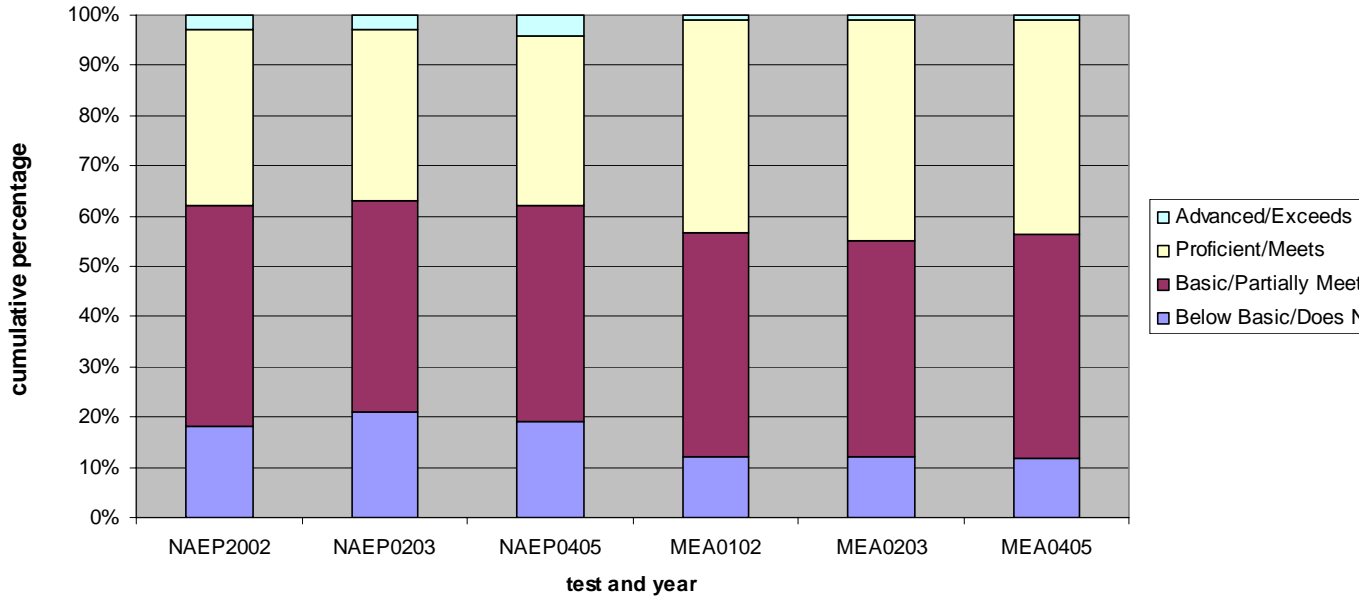
Appendix A – Comparison of MEA and NAEP Achievement Level Trend Lines

Note: More appendices will be attached analysis continues.

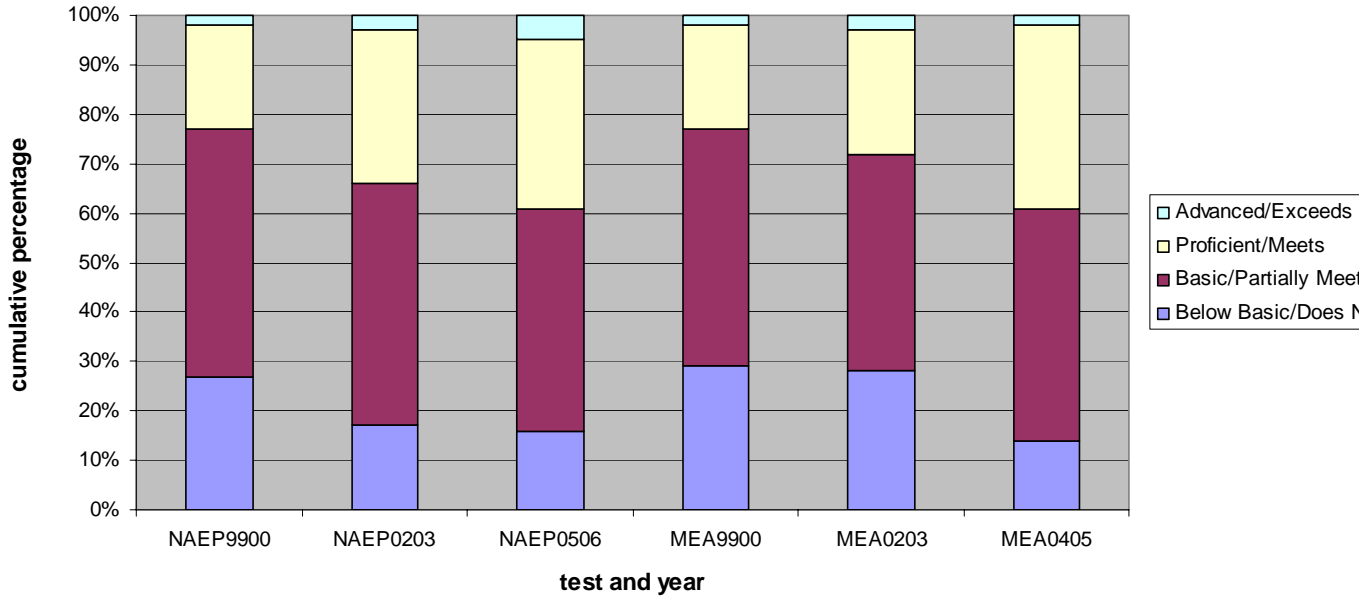
Percentages of Students at Performance Levels for Grade 4 Reading NAEP and MEA Tests



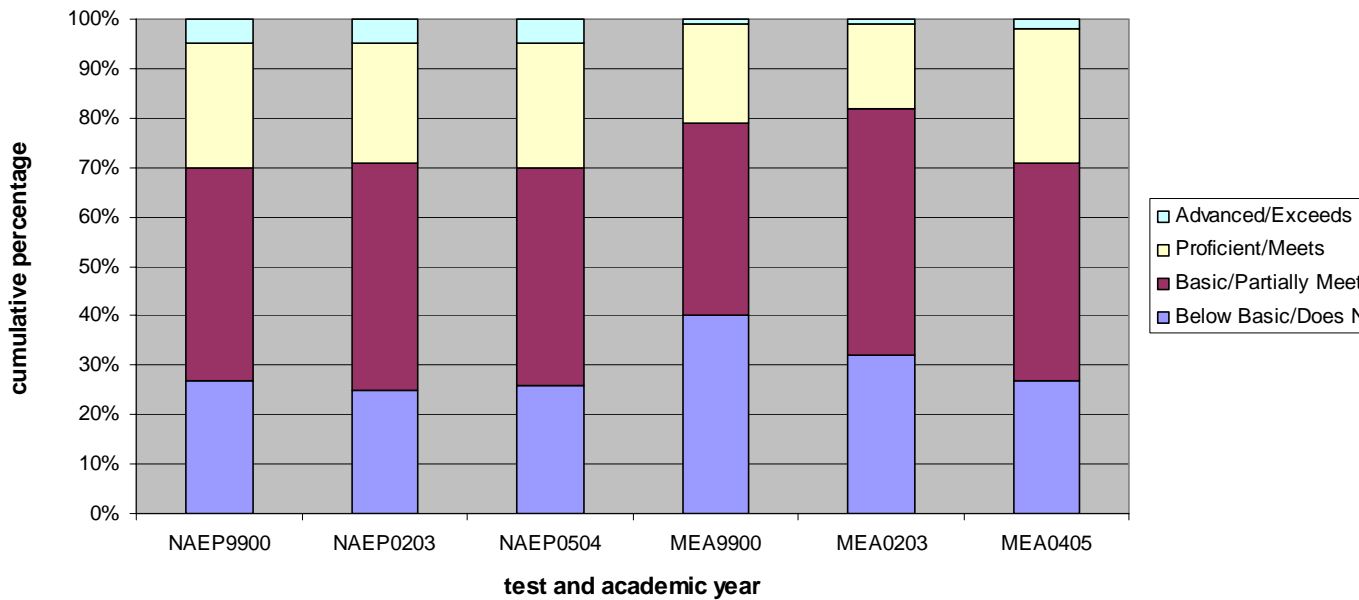
Percentages of Students at Performance Levels for Grade 8 Reading NAEP and MEA Tests



Percentages of Students at Performance Levels for Grade 4 Mathematics NAEP and MEA Tests



Percentages of Students at Performance Level for Grade 8 Mathematics NAEP and MEA Tests



Tables prepared by Tad Johnston, MEDOE

4 p.m. Tuesday, October 18, 2005