

Report on 2007
Trial Urban District Assessment (TUDA)
National Assessment of Educational
Progress (NAEP)

Grade 8 Writing

THE SCHOOL COMMITTEE OF THE CITY OF BOSTON

Elizabeth Reilinger, Chair

Rev. Gregory G. Groover, Sr., Vice-Chair

Michele P. Brooks, Member

Helen M. Dájer, Member

Alfreda Harris, Member

Claudio Martinez, Member

Marchelle Raynor, Member

SUPERINTENDENT OF SCHOOLS

Carol R. Johnson

OFFICE OF RESEARCH, ASSESSMENT, AND EVALUATION

Ru-Ing Hwang, Acting Director

TABLE OF CONTENTS

Executive Summary.....	i
Overview and Background.....	1
Demographic Context.....	2
Analyses.....	3
• Comparisons of Average Scale Score: Boston and TUDA Districts	4
• Comparisons of Students At or Above Proficiency Level: Boston and TUDA Districts.....	5
• Comparisons of Performance by Selected Race/Ethnicity and Score Gaps Between Selected Racial/Ethnic Groups: Boston and TUDA District	6
• Comparisons of Performance by Various Student Subgroups: Boston and TUDA Districts.....	10
• Comparisons of Average Scale Score by Various Student Groups: Boston and Massachusetts.....	12
Appendix A: Assessment Framework	
Appendix B: Comparison of NAEP and MCAS	
Appendix C: Samples of Released 2007 NAEP Writing Prompts	
Appendix D: Average Scale Scores and Achievement-Level Results by Selected Race/Ethnicity and Jurisdiction	
Appendix E: Average Scale Scores and Achievement-Level Results by Various Student Groups and Jurisdiction	

EXECUTIVE SUMMARY

The Trial Urban District Assessment (TUDA) was started in 2002 as a part of the National Assessment of Educational Progress (NAEP). The Boston Public Schools was one of eleven urban districts that voluntarily participated in the NAEP assessment on a trial basis. In 2007, Boston participated in grade 8 writing in addition to reading and mathematics in grades 4 and 8. Results on NAEP 2007 grade 8 writing for Boston are compared with other TUDA districts, Large Central Cities (LCC), and Massachusetts.

Overall 2007 Findings

- **LCC's:** Boston's 2007 overall writing performance at grade 8 is not significantly different from that in LCC in terms of average scale score and percentage of students performing at or above the Proficient level (Boston: 149, 25%; LCC: 145, 22%, respectively).
- **Other TUDA Districts:** Boston's grade 8 writing performance is higher¹ than those in three urban districts (Houston, Los Angeles, and Cleveland) in terms of average scale score and percent of students performing at or above the Proficient level. Boston is exceeded by only one district (Charlotte) in terms of average scale score (155), but no difference is found in terms of percent of students performing at or above Proficient (31%) compared with other TUDA districts.

Performance by Selected Racial/Ethnic Groups

- **LCC's:** White students in Boston outperformed similar students in large central cities in terms of average scale score and percentage of students performing at or above Proficient.
- **Other TUDA Districts:**
 - **Average Scale Score:** the average scale scores for Black and White students in Boston are higher than those in Cleveland and Los Angeles and are about the same as the other districts except for Black students in Boston that also scored higher than that in Austin. Hispanic students in Boston on average scored higher than their counterparts in San Diego, but performed lower than those in Chicago.
 - **Percentage At or Above Proficient:** the percent of Black and White students in Boston performing at or above Proficient is greater than similar students in Cleveland.
- **Score Gaps:**
 - **White – Black:** Boston's average scale score gap is statistically no different than any other district or LCC except for Cleveland which has a smaller gap than Boston.

¹ Any difference noted as "higher" or "lower" in this report denotes a statistically significant difference.

- **White – Hispanic:** Boston's average scale score gap is statistically no different than any other district or LCC except for Cleveland which has a smaller gap.
- It should be noted that White students in Cleveland scored 31 points lower than those in Boston.
- The gaps in performance between Whites and Blacks/Hispanics are seen in urban districts across the nation.

Performance by Various Student Groups

- Boston's female students outperformed male students.
- **LCC's:** Boston's female students, students with a disability, and low-income students on average scored higher than similar students in LCC, but the scores for English language learners (ELL) in Boston are lower than that in LCC. However, the percent of students performing at or above the Proficient level for Boston is about the same as that in LCC for all student groups.
- **Other TUDA Districts:** Boston consistently scored higher than Cleveland and Los Angeles for the following student groups: Female, Male, Students with a disability, and eligible for free/reduced lunch. Students in Boston with a disability also outperformed their peers in Chicago, and the score for Boston's low-income students also exceed those in Austin and San Diego. The scores for all student groups in the other districts are comparable to the corresponding scores for Boston except for Charlotte and Los Angeles. Female students in Charlotte and ELL students in both Charlotte and Los Angeles had a higher score than similar students in Boston.
- A similar pattern in terms of percent of students performing at or above Proficient is seen for those TUDA districts that met the reporting standard.

Comparisons between Boston and Massachusetts

- Boston's average scale score is 18 points lower than the state.
- Among the racial/ethnic groups, only Black students in Boston had a score that is lower than their peers across the state.
- Both female and male students in Boston on average scored 18 points lower than their counterparts statewide.
- In Boston, the scores for students with disabilities are lower than similar students across the state by 18 points.
- It should be noted that although Boston has a higher percentage of ELL and low-income students, the average scale scores for these two student groups are no different from similar students statewide.

OVERVIEW AND BACKGROUND

In 2001, after discussion among the National Center for Education Statistics (NCES), the National Assessment Governing Board (NAGB), and the Council of the Great City Schools (CGCS), Congress appropriated funds for a district-level assessment on a trial basis, similar to the trial for state assessments that began in 1990. The NAGB passed a resolution approving the selection of urban districts for participation in the Trial Urban District Assessment (TUDA), a special project within NAEP. Representatives of the Council of Great City Schools worked with the staff of NAGB to identify districts to be invited for the trial assessment. Districts were selected based on a range of characteristics, such as district size, minority concentrations, federal program participation, socioeconomic conditions, and percentages of students with disabilities (SD) and English Language Learners (ELL) students.

This report provides selected results from the National Assessment of Educational Progress (NAEP) for public school students at grade 8. While national and state-level writing results are available for three assessment years (1998, 2002, and 2007), district-level results in writing are available for 2002 and 2007 for only four of the TUDA districts. Since Boston joined the TUDA in 2003, there is no comparative information from 2002 to 2007 for Boston.

In the 2007 assessment, Boston was one of 10 urban school districts that participated and met the criteria for reporting public school results. The other districts were: Atlanta, Austin, Charlotte-Mecklenburg, Chicago, Cleveland, Houston, Los Angeles, New York City, and San Diego.

In Boston, grade 8 students from 34 middle schools were sampled. These students are also included in “higher-level” samples. For example, data from students tested in the Boston sample were used to report results for Boston, and also contributed to the Massachusetts and the national samples.

In the report tables and graphs, Boston results are compared with the results from other TUDA participating districts or public school students in the in large central cities. “Large Central City” (LCC) in this report includes public schools located in large central cities (population of 250,000 or more as defined by NCES). Three of the 10 TUDA districts (Austin, Charlotte and Los Angeles) consist primarily of students enrolled in schools in large central cities, but also have a number of their eighth grade students enrolled in surrounding suburban or rural areas. It should be noted that the “Large Central City” is not synonymous with “inner city.”

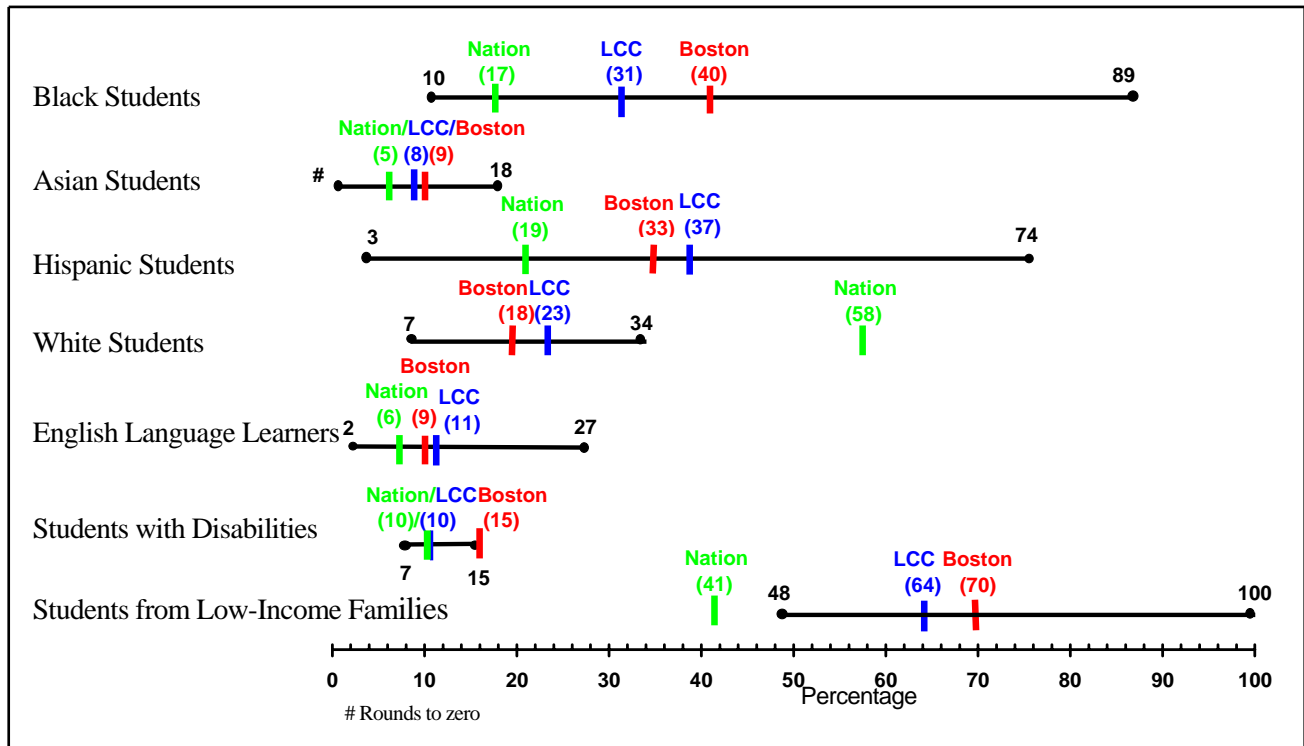
An overview of the assessment framework and comparisons with the MCAS relative to design, reporting and format are included in Appendices A and B. Appendix C presents samples of released 2007 NAEP grade 8 writing prompts.

DEMOGRAPHIC CONTEXT

The graphic on below displays the percentages of subgroups (Black, Hispanic, English Language Learner, Students with Disabilities, Students from Low-Income Families) for the nation, for Boston Public Schools, for LCC and the range for the TUDA districts. The percentages are based on grade 8 students who participated in the 2007 TUDA NAEP Writing test.

For every racial/ethnic group and the eligibility for free/reduced-price lunch group, Boston's percentage is in the middle range of the other TUDA districts, while the percentage of ELL population is slightly lower. Of note is that Boston has the highest percentage of students with disabilities that participated in the test. These aspects are important to keep in mind when reviewing the results.

Range of Percentage of Students in Each Student Group Who Participated in the 2007 Grade 8 NAEP Writing Assessment for TUDA Districts



ANALYSES

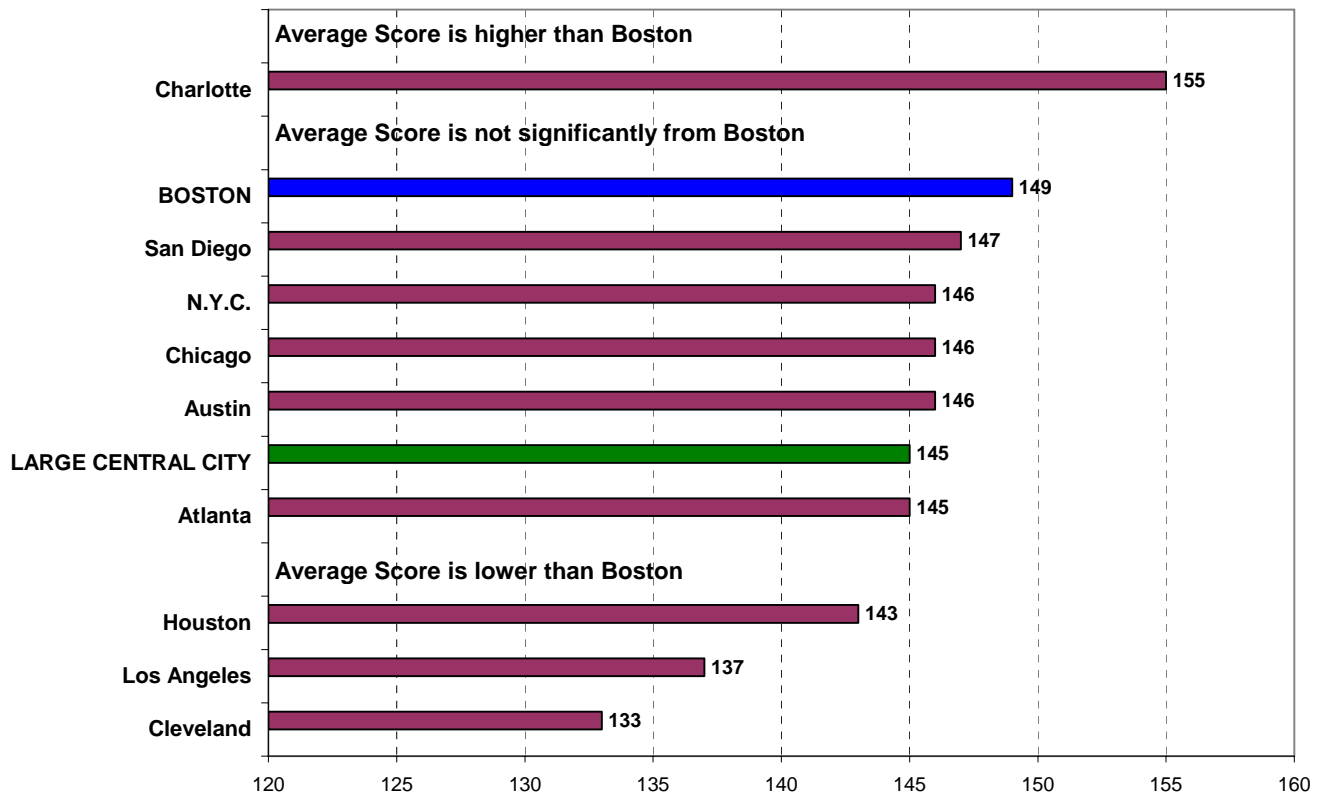
Performance is examined in five ways:

- (1) comparisons of average scale scores with other communities that participated in the TUDA project and the large central city schools. This provides normative information relative to specific other large cities.
- (2) percent of students performing at or above Proficiency. This provides information on the top level of student performance between Boston and other TUDA districts. Given that NCLB requires that all students must reach proficiency by 2014, it is useful to examine performance at this level.
- (3) performance of selected racial/ethnic groups and score gaps between selected racial/ethnic groups in Boston compared to that of other TUDA districts and in large central cities. This provides information on achievement levels for various racial/ethnic subgroups.
- (4) comparative performance of students of low-income backgrounds, English language learners, students with disabilities, and by gender. This provides normative information relative to specific other large cities for various student subgroups.
- (5) performance of Boston compared to Massachusetts (MA) statewide, by various students groups. This provides information relative to performance of similar students statewide.

Because the NAEP results are based on a sample of students within a district (TUDA) or across Massachusetts and not the entire population of the jurisdiction, examining statistical significance is critical to determine differences across groups.

(1) Scale Score Comparisons of 2007 NAEP Grade 8 Writing Between Boston, Large Central Cities, and Other Participating TUDA Districts

Boston’s Average Scale Score on 2007 NAEP Grade 8 Writing Compared with Scores for the LCC and Other Participating Jurisdictions



Note: The NAEP Writing scale ranges from 0 to 300.
 The bars are graphed using rounded numbers. Significance tests used a multiple-comparison procedure based on unrounded numbers from all jurisdictions that participated.

Comparison with Large Central Cities

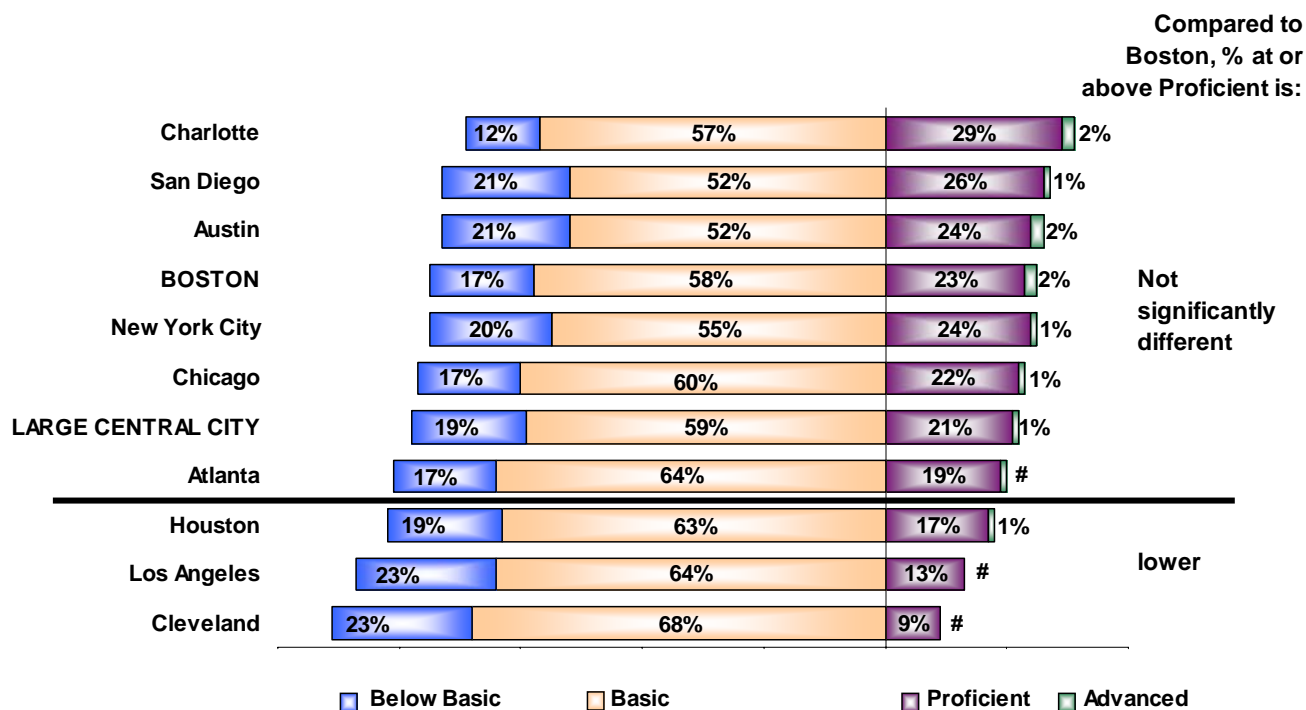
- Boston’s grade 8 writing average scale score (149) is not significantly different from that for the large central city schools (145).

Comparisons with Other TUDA Districts

- Performance in Boston in grade 8 writing is significantly higher than the scores in three urban districts (Houston, Los Angeles and Cleveland), not significantly different from those in 5 districts (San Diego, New York, Chicago, Austin, and Atlanta) and lower than that in Charlotte.

(2) Comparisons of Percentage of Students Performing At or Above Proficient on 2007 NAEP Grade 8 Writing Between Boston, Large Central Cities, and Other Participating TUDA Districts

Boston's Percentage At or Above Proficient Level in 2007 NAEP Grade 8 Writing Compared with Percentage for the LCC and Other Participating Jurisdictions



Note: # Rounds to zero.

The bars are graphed using rounded numbers. Significance tests used a multiple-comparison procedure based on unrounded numbers from all jurisdictions that participated. Percents may not sum to 100% because of rounding.

Comparison with Large Central Cities

- The percentage of students performing at or above proficient level in Boston (25%) is not significantly different from that in large central cities (22%).

Comparisons with Other TUDA Districts

- While the percentage of students in Boston who performed at or above proficient level (25%) is significantly higher than those in Houston (18%), Los Angeles (13%) and Cleveland (9%), Boston scored about the same as their peers in the other six participating urban districts: Charlotte (31%), San Diego (27%), Austin (26%), New York City (25%), Chicago (23%), and Atlanta (19%).

(3) Comparisons of Performance by Race/Ethnicity on 2007 NAEP Grade 8 Writing Between Boston, Large Central Cities, and Other Participating TUDA Districts

Boston's Average Scale Scores and Percentage At or Above the Proficient Level in 2007 NAEP Grade 8 Writing Compared with that of the LCC and Other Participating Jurisdictions: By Race/Ethnicity

<i>Race / Ethnicity (Boston Score)</i>	LCC*	Atlanta	Austin	Charlotte	Chicago	Cleveland	Houston	Los Angeles	N.Y.C.	San Diego
Average Scale Score										
African American/Black (141)	=	=	↑	=	=	↑	=	↑	=	=
Asian/Pacific Islander (174)	=	⊘	⊘	⊘	⊘	⊘	=	=	=	=
Hispanic (138)	=	⊘	=	=	↓	=	=	=	=	↑
White (173)	↑	=	=	=	=	↑	=	↑	=	=
Percent At or Above Proficient										
African American/Black (16%)	=	=	=	=	=	↑	=	=	=	=
Asian/Pacific Islander (55%)	=	⊘	⊘	⊘	⊘	⊘	=	=	=	=
Hispanic (14%)	=	⊘	=	=	=	=	=	=	=	=
White (52%)	↑	=	=	=	=	↑	=	=	=	=

* LCC: Large Central Cities

Relative to each district listed at the top of the figure:

↑ : Boston had significantly higher average scale score or percent at or above Proficient

= : No significant difference between scale scores or percent at or above Proficient

↓ : Boston had significantly lower average scale score or percent at or above Proficient

⊘ : Reporting standards not met

Note: (1) Significance tests between Boston and each jurisdiction used a multiple-comparison procedure based on all jurisdictions that participated.
 (2) Source for the significant testing on performance among racial/ethnic groups within Boston is "U.S. Department of education, Institute of Education Science, National Center for Education Statistics, National Assessment of Education Progress (NAEP), 2007 Trial Urban District Writing Assessment.

By Average Scale Score

Within Boston:

- White students in Boston had an average scale score that was higher than those of Black and Hispanic students, but was not found to be significantly different from that of Asian/Pacific Islander students.

Compared to LCC and Other TUDA Districts:

- Compared to the performance of students in large central cities, White students in Boston had a significantly higher average scale score. The other racial/ethnic groups (Black, Hispanic and Asian) scored about the same as those in large central cities.
- The average scale scores of White and Black students in Boston are higher than in Cleveland and Los Angeles and are not significantly different from those in other districts, except for Black students in Boston, who also scored significantly higher than their peers in Austin.
- While the average scale score of Hispanic students in Boston is significantly higher than that in San Diego, Boston's Hispanic students scored significantly lower than their counterparts in Chicago and are similar to Hispanic students in the rest of TUDA districts.
- The average scale score of Asian students in Boston are not significantly different from the scores for the other TUDA districts that met the reporting standards.

By Achievement Level: Percent At or Above Proficient

Within Boston:

- In Boston, the percentage of White students performing at or above Proficient is greater than those of Black and Hispanic students, but is not significantly different between White and Asian students.

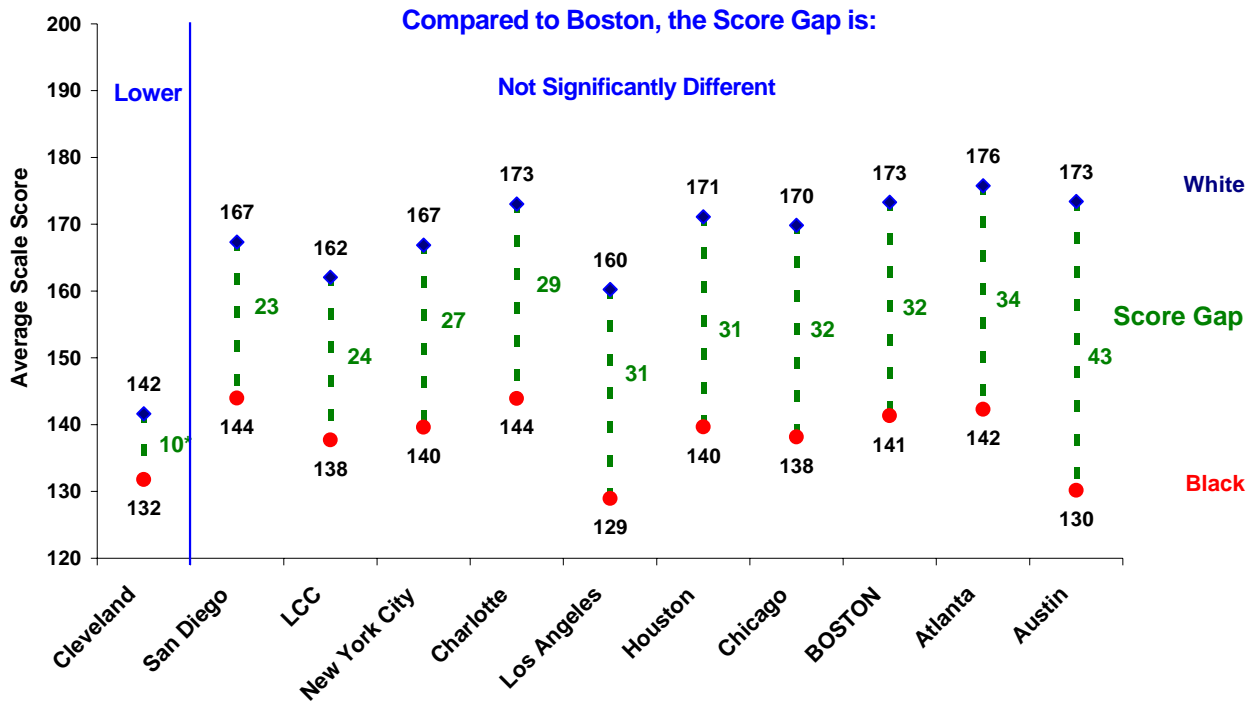
Compared to LCC and Other TUDA Districts:

- The percentage of White students in Boston performing at or above Proficient exceeds that in large central cities, but no significant difference is found for the other racial/ethnic groups.
- The respective percentages of White and Black students in Boston performing at or above Proficient are comparable to all participating districts except for Cleveland which had a lower percentage than Boston.
- The percentage of Asian and Hispanic students in Boston scored at or above Proficient level is not significantly different from all other TUDA districts that met the sample size requirements.

By Score Gaps

Comparisons of average scale score gaps between White and Black or White and Hispanic students in Boston, TUDA districts, and LCC are presented in the figures shown below.

**2007 NAEP Grade 8 Writing
White - Black Average Scale Score Gaps**

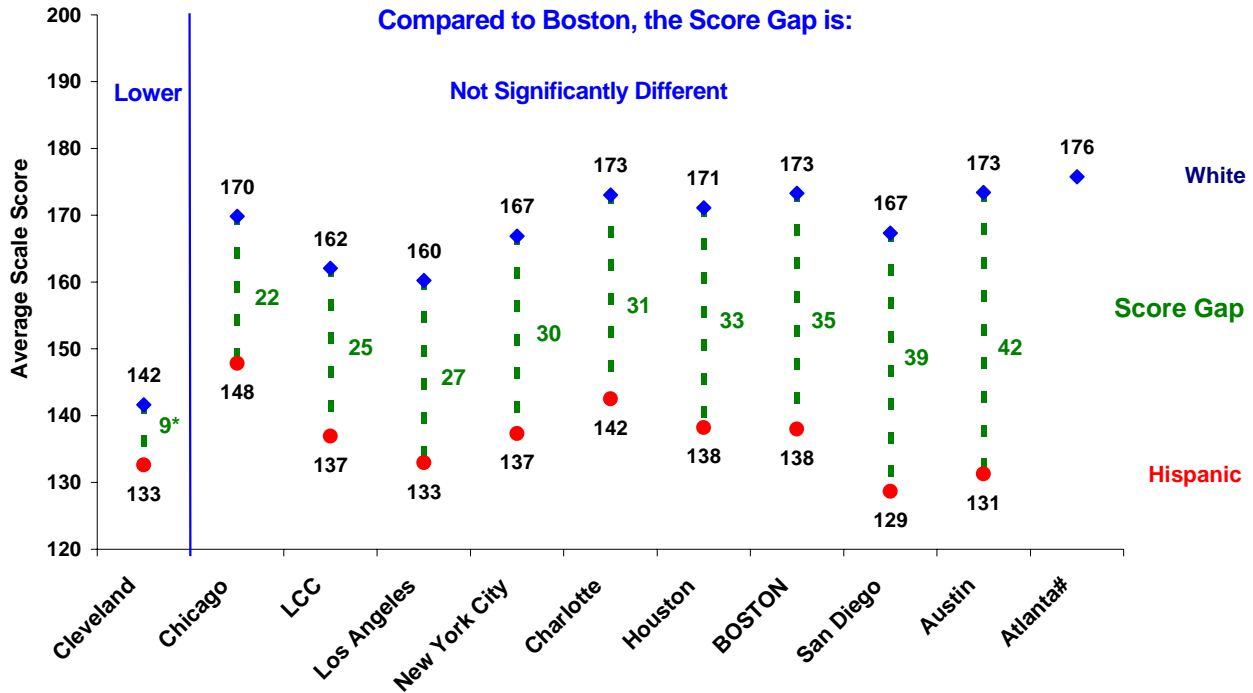


Note: The NAEP Writing scale ranges from 0 to 300.

The bars are graphed using unrounded numbers. Significance tests used a multiple-comparison procedure based on all jurisdictions that participated.

- The average scale score gap between White and Black students ranges from a low of 10 points in Cleveland to a high of 43 points in Austin. However, the score gap is not significantly different between Boston (32 points) and all other districts or large central cities except for Cleveland which has a smaller gap.
- It should be noted that both White and Black students in Cleveland scored lower than similar students in Boston: 30 points lower for White and 9 points for Black students.
- Clearly, the gaps in writing performance between White and Black students exist not only in Boston but also for TUDA districts and LCC.

**2007 NAEP Grade 8 Writing
White - Hispanic Average Scale Score Gaps**



Note: The bars are graphed using unrounded numbers. Significance tests used a multiple-comparison procedure based on all jurisdictions that participated.

Reporting standards not met for Hispanic students. Sample size is insufficient to permit a reliable estimate.

- Similar score gap pattern can be seen in the differences of average scale scores between White and Hispanic students across all jurisdictions that ranges from a low of 9 points in Cleveland to a high of 42 points in Austin. Only Cleveland had a smaller gap than that in Boston (35 points). The scale score gaps for all other districts are similar to that of Boston.
- It should be noted that both White and Hispanic students in Cleveland scored lower than those in Boston: 31 points lower for White and 5 points for Hispanic students.
- The gap in writing scores between White and Hispanic students is clearly a challenging issue faced by large school districts across the nation.

Average scale score and achievement-level for each districts and large central cities by selected race/ethnicity categories may be found in Appendix D.

(4) Comparisons of Performance by Other Student Groups on 2007 NAEP Grade 8 Writing Between Boston, Large Central Cities, and Other Participating TUDA Districts

Boston's Average Scale Scores and Percentage At or Above the Proficient Level in 2007 NAEP Grade 8 Writing Compared with that of the LCC and Other Participating Jurisdictions: By Various Student Subgroups

<i>Student Subgroups (Boston Score)</i>	LCC*	Atlanta	Austin	Charlotte	Chicago	Cleveland	Houston	Los Angeles	N.Y.C.	San Diego
Average Scale Score										
Female (160)	↑	=	=	↓	=	↑	↑	↑	=	=
Male (138)	=	=	=	=	=	↑	=	↑	=	=
Students with Disabilities (121)	↑	=	=	=	↑	↑	=	↑	=	=
English Language Learners (102)	↓	⊘	=	↓	=	⊘	=	↓	=	=
Eligible Free/Reduced-Price Lunch (144)	↑	=	↑	=	=	↑	=	↑	=	↑
Percent At or Above Proficient										
Female (35%)	=	=	=	=	=	↑	↑	↑	=	=
Male (15%)	=	=	=	=	=	↑	=	↑	=	=
Students with Disabilities (4%)	=	⊘	=	=	=	⊘	=	=	⊘	=
English Language Learners (1%)	⊘	⊘	⊘	⊘	⊘	⊘	⊘	⊘	⊘	⊘
Eligible Free/Reduced-Price Lunch (18%)	=	=	↑	=	=	↑	=	↑	=	=

* LCC: Large Central Cities

Relative to each district listed at the top of the figure:

- ↑ : Boston had significantly higher average scale score or percent at or above Proficient
- = : No significant difference between scale scores or percent at or above Proficient
- ↓ : Boston had significantly lower average scale score or percent at or above Proficient
- ⊘ : Reporting standards not met

Note: (1) Significance tests between Boston and each jurisdiction used a multiple-comparison procedure based on all jurisdictions that participated.
 (2) Source for the significant testing on performance between gender groups within Boston is "U.S. Department of education, Institute of Education Science, National Center for Education Statistics, National Assessment of Education Progress (NAEP), 2007 Trial Urban District Writing Assessment.

By Average Scale Score

Within Boston:

- Male students in Boston had an average scale score (138) that was lower than the score for female students (160).

Compared to LCC and Other TUDA Districts:

- Female students in Boston scored significantly higher than that in large central cities, Cleveland, Houston, and Los Angeles but scored lower than Charlotte, and scored about the same as the other five districts.
- Male students in Boston on average scored about the same as their peers in all other TUDA districts or large central cities except for Cleveland and Los Angeles where male students had lower scores.
- The average scale score for students with disabilities in Boston is significantly higher than scores for their counterparts in Chicago, Cleveland, Los Angeles and large central cities and comparable to scores for the other six TUDA districts.
- English language learners in Boston on average scored significantly lower than those in large central cities, Charlotte, and Los Angeles and scored about the same compared with the other districts that met the reporting standards.
- The average scale score for low-income students (as measured by eligibility for the Free/Reduced Price National School Lunch Program) in Boston is higher than the scores in four districts (Austin, Cleveland, Los Angeles and San Diego) and large central cities, and not significantly different than five districts.

By Achievement Level: Percent At or Above Proficient

Within Boston:

- A significant higher percentage of female than male students in Boston performed at or above Proficient level, 35% and 15%, respectively.

Compared to LCC and Other TUDA Districts:

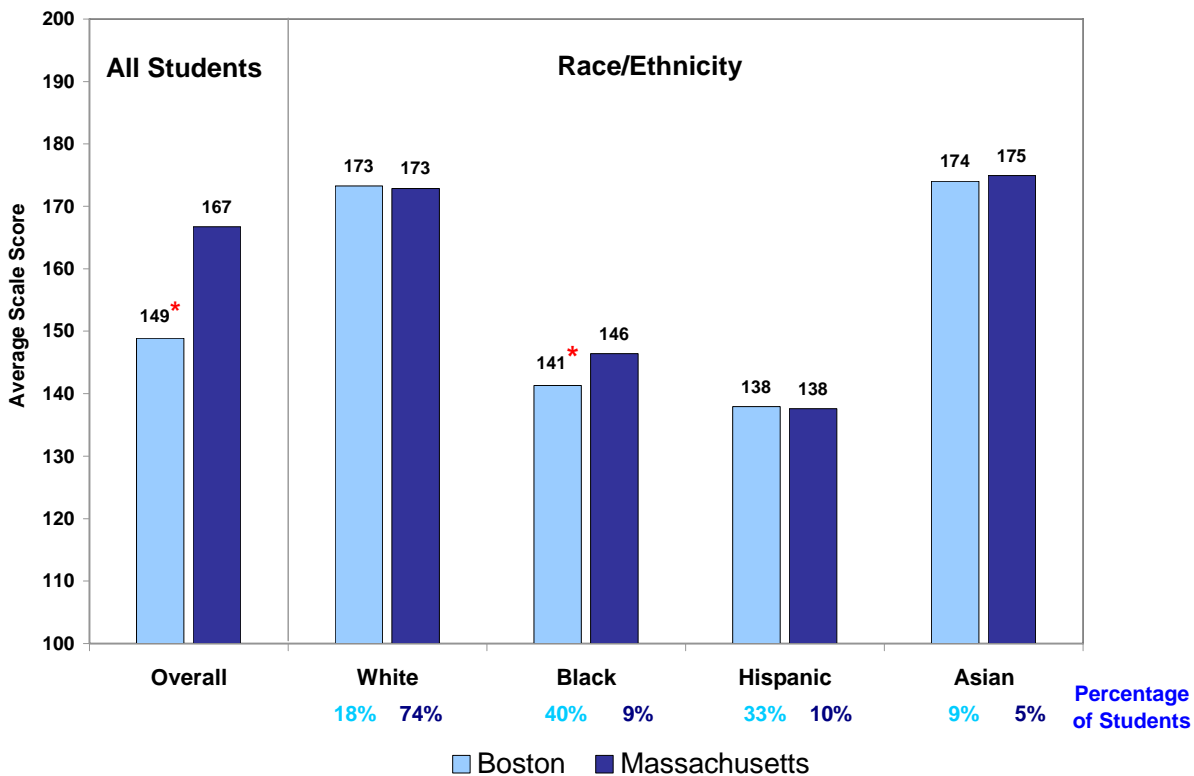
- The percentage of female students performing at or above Proficient level in Boston is higher than the percentages in Cleveland, Houston and Los Angeles, and are comparable to those in five other TUDA districts and large central cities.
- For male students in Boston, the percentage at or above Proficient was not significantly different from all participating jurisdictions except for Cleveland and Los Angeles which had lower percentages.
- The percentage of students with disabilities performing at or above Proficient in Boston is comparable to similar students in large central cities and across all TUDA districts that met the reporting standards.
- No participating district had a sufficiently large enough group of English language learners to permit a reliable estimate.
- There is no significant difference in the percentage of low-income students performing at or above Proficient between Boston and similar students in all other

TUDA districts and large central cities except for Austin, Cleveland and Los Angeles that had lower percentages.

Full information is contained in Appendix E.

(5) Scale Score Comparisons of 2007 NAEP Grade 8 Writing Between Boston and Massachusetts statewide, by various students groups

Boston's Average Scale Scores in 2007 NAEP Grade 8 Writing Compared with Scores for the Massachusetts: Overall and by Race/Ethnicity

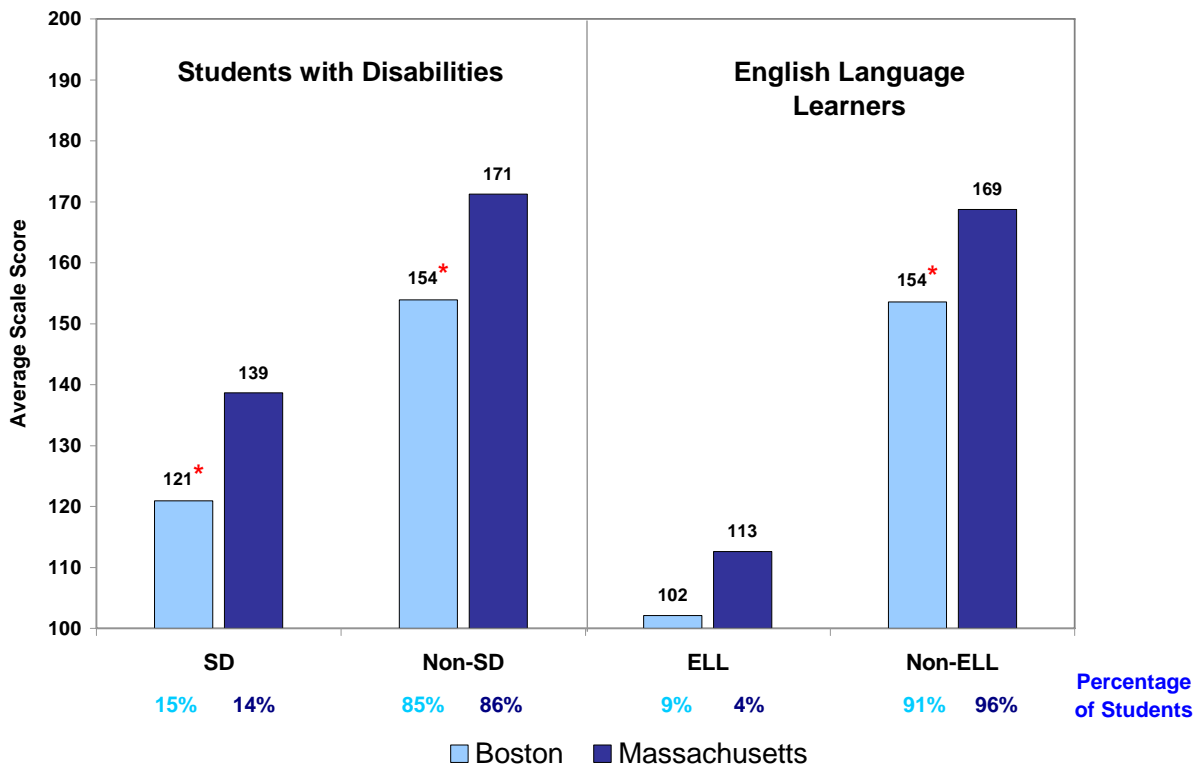


Note: The NAEP Writing scale ranges from 0 to 300

* Boston had significantly lower average scale score than the state.

- Overall, Boston 8th graders scored 18 points lower than their peers statewide on 2007 NAEP writing.
- The average scale score for all racial/ethnic groups in Boston are comparable to their peers statewide except for Black students that scored about 5 points lower.
- It should be noted that Boston has about 4.5 times more of Black students (40%) than similar students statewide (9%).

Boston's Average Scale Scores in 2007 NAEP Grade 8 Writing Compared with Scores for the Massachusetts: By Students with Disabilities and By English Language Learners

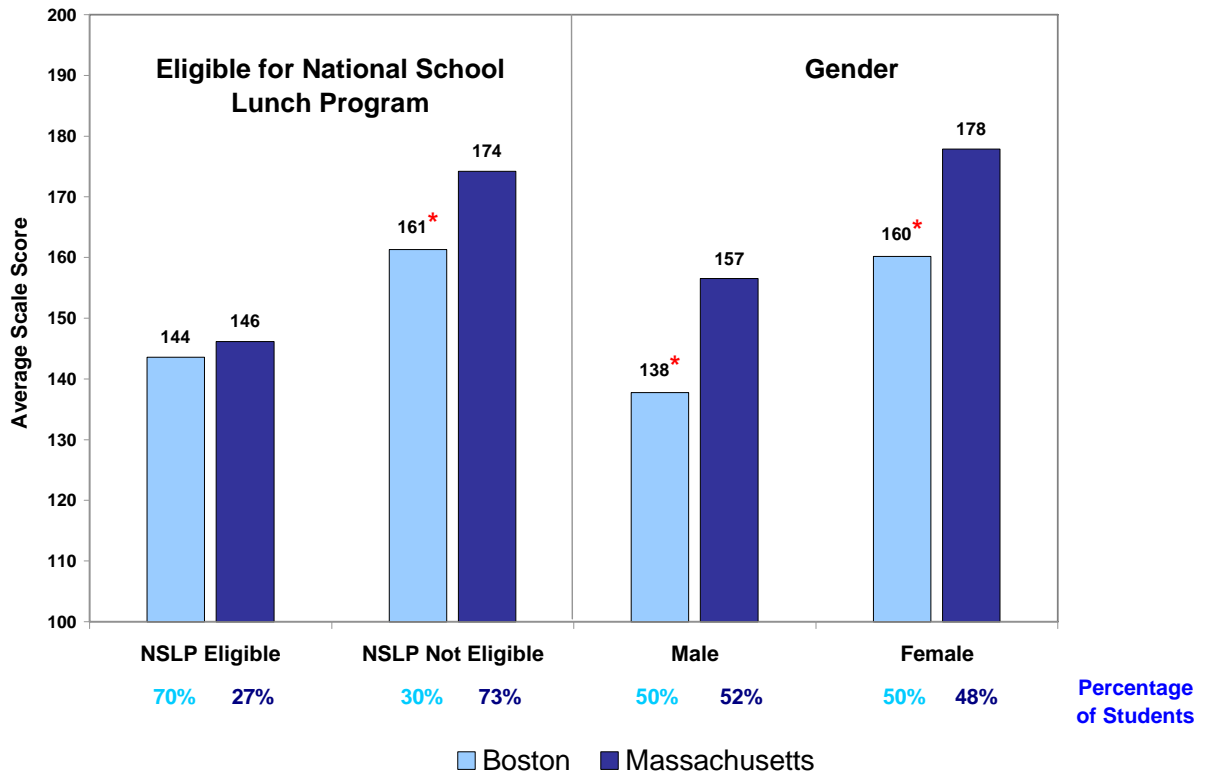


Note: The NAEP Writing scale ranges from 0 to 300

* Boston had significantly lower average scale score than the state.

- Both students with or without a disability in Boston scored lower than their counterparts across the state by 18 points and 17 points, respectively.
- While the average scale score of English language learners in Boston is not significantly different from that for similar students statewide, Boston's Non-ELL students scored 15 points lower than their peers from the state.

Boston's Average Scale Scores in 2007 NAEP Grade 8 Writing Compared with Scores for the Massachusetts: By School Lunch Eligibility and By Gender



Note: The NAEP Writing scale ranges from 0 to 300

* Boston had significantly lower average scale score than the state.

- While Boston has substantially more students (70%) who are eligible for free/reduced lunch program than statewide (27%), the average scale score of low-income students in Boston is statistically not different from similar students statewide. However, Boston's non-eligible students scored 13 points lower than their peers from the state.
- Both male and female students in Boston had an average scale score that is 19 points and 18 points lower than their counterparts across the state, respectively.

NEXT STEPS

Despite the fact that Writing is not a separate subject area that is being held accountable under the No Child Left Behind Act, it is one of the most important forms of communication. Good writers can communicate well in a range of situations. The ability to write well requires organizing thoughts, analyzing arguments and synthesizing information. To improve writing skills students need expert instruction, frequent practice, and constructive feedback. Hence, it is suggested that:

- In conjunction with the office of Teaching and Learning, this writing results along with MCAS writing prompt results can be used to assess strengths and weaknesses based on student performance on the released student responses.
- Boston should continue the implementation of professional development and writing process/writers workshop in the district.

APPENDIX A: Assessment Framework

Writing

The NAEP writing framework, which defines the content for the 2007 assessment, was developed through a comprehensive national consultative process and adopted by NAGB. The writing framework emphasizes that the NAEP writing assessment should have the following objectives:

- Students should write for a variety of purposes: *narrative-telling a story*, *informative-informing the reader*, and *persuasive-persuading the reader*;
- Students should write on a variety of tasks and for many different audiences;
- Students should write from a variety of stimulus materials, and within different time constraints;
- Students should generate, draft, revise, and edit ideas and forms of expression in their writing;
- Students should display effective choices in the organizations of their writing. They should include detail to illustrate and elaborate their ideas, and should use appropriate conventions of written English; and
- Students value writing as a communicative activity.

The writing framework also specifies the distribution of tasks by grade and writing purpose to reflect differing levels of student development and instructional focus. The percentage distribution of questions for grade 8 is as follows: Narrative: 33%, Informative: 33%, and persuasive: 33%. Each student takes two (out of 17 tasks) 25-minute writing tasks.

A new framework is under development for the 2011 writing assessment.

Accommodations

It is NAEP's intent to assess all selected students from the target population. Beginning in 2002, students with disabilities and limited-English-proficient students who require accommodations have been permitted to use them in NAEP, unless a particular accommodation would alter the skills and knowledge being assessed. For example, NAEP does not permit the test materials to be read aloud in native languages.

Population Tested

Results from the 2007 Trial Urban District Assessment are reported for the participating districts for public-school students at grade 8. The TUDA employed larger-than-usual samples within the districts, making reliable district-level data possible. The samples were also large enough to provide reliable estimates on subgroups within the districts, such as female students or Hispanic students. Because students were sampled, all analyses are examined for significant significance. Grade 8 Writing was last done nationally, statewide and for TUDA districts in 2002, so there are no longitudinal data for Boston.

In Boston, 1,113 students from 34 schools at grade 8 participated in the 2007 NAEP assessments.

(Intentionally left blank)

Appendix B



NAEP vs. MCAS

Introduction

Under the federal *No Child Left Behind Law* (NCLB) and state *Education Reform Law of 1993*, Boston Public School students are required to participate in two testing programs: the National Assessment for Educational Progress (NAEP) and the Massachusetts Comprehensive Assessment System (MCAS). The biennial NAEP Trial Urban School District Assessment (TUDA) provides important information for understanding the effective of BPS system relative to other large urban school districts, while the annual MCAS test provides critical information about the academic performance of BPS compared to other Mass. Public schools as well as to what extent BPS students achieve, on the Mass. Curriculum standards.

This section will briefly compare MCAS with NAEP, and provide information on interpreting the test results and how to make the comparisons and/or connections between the two assessments.

Overview

NAEP

- The National Assessment of Educational Progress (NAEP), known as the Nation's Report Card, is a Congressionally-mandated assessment since 1969. It includes state assessment since 1990 and conducted the first Trial Urban School District Assessment (TUDA) in 2002. Based on policy set by the National Assessment Governing Board (NAGB), NAEP measures what students know and can do in key subject areas.

MCAS

- The Massachusetts Comprehensive Assessment System (MCAS), fulfilling the requirements of the Education Reform Act of 1993, is the Commonwealth's statewide assessment program for public schools since 1998.

Requirements for Student Participation

Student Selection

NAEP

- Based on sampling, a representative sample from randomly selected schools must participate in NAEP testing. For Trial District Assessment, the target sample size per subject per grade is 1200-1400 students. About 30 students per subject, at each participating school are tested.

MCAS

- All Massachusetts public school students in the grades tested must take the MCAS tests.



Student Participation

NAEP

- Beginning in 2003, school receiving Title I subgrants are required to participate in the biennial NAEP assessments. In 2007, students participated in the writing assessment at grades 8 & 12 (national only) if selected for the NAEP sample. Under NCLB, parental notification prior to testing is mandatory to inform parents of students who are sampled that their child's participation is voluntary.

MCAS

- The writing assessment is part of the ELA MCAS assessment in grades 4, 7, and 10. Every public school student in grades 3 thru 8 & 10 is mandated to take the ELA test.

Inclusions & Accommodations

NAEP

Includes students with disabilities and English Language Learners (ELL) students in the assessment.

- **ELL:** NAEP includes all ELL students who have received instruction in English for at least three years. ELL students who have received instruction in English for less than three years are included as well unless school staff judge them to be incapable of participating in the assessment in English.
- **Students with Disabilities (SD):** Based on student's IEP/504 plan, students with disabilities are tested with appropriate accommodations unless the student's IEP team judges that he or she cannot participate or if NAEP does not allow an accommodation that the student requires.
- In general, inclusion in NAEP of an SD or ELL student is encouraged if that student (a) participated in the regular state academic assessment in the subject being tested, and (b) if that student can participate in NAEP with the accommodations NAEP allows.

MCAS

Includes students with disabilities and limited English Proficient (LEP) students in the assessment.

- **LEP:** Beginning in 2003, *No Child Left Behind* as well as *Question 2*, the Massachusetts ballot initiative approved by voters in November 2002, requires that all LEP students participate in state administered academic assessments, with the sole exception of LEP students in their first year of enrollment in U.S. schools. Schools have the option of administering the ELA and History/Social Science tests to first-year LEP students.
- **Students with Disabilities:** The vast majority of students with disabilities take standard MCAS tests, either with or without accommodations as specified in their IEP/504 plan. Only a very small number of students with the most significant disabilities take the MCAS Alternate Assessment.



Test Content/Instrument Design

Framework

NAEP

The content and design of the NAEP writing assessment was constructed based on the Frameworks that were developed by the National Assessment Governing Board (NAGB).

- **Writing:** The 1998 updated NAEP Writing Framework

MCAS

The content knowledge and skills tested by MCAS were based on the learning standards in the Massachusetts Curriculum Framework for these content areas.

- **English Language Arts:** Massachusetts English Language Arts Curriculum Framework, June 2001 and May 2004 Supplement

Content Standards Tested

NAEP

Writing: assesses three contexts for writing

- **Narrative writing:** Narrative writing encourages writers to incorporate their imagination and creativity in the production of stories or personal essays. At its best, narrative writing fosters imagination, creativity, and speculation by allowing writers to express their thoughts and emotions, and offers an opportunity for writers to analyze and understand their actions and those of others.
- **Informative writing:** In informative writing, the writer provides the reader with information. This type of writing is used to share knowledge and to convey messages, instructions, and ideas. When used as a means of exploration, informative writing helps both the writer and the reader to learn new ideas and to reexamine old conclusions.
- **Persuasive writing:** Persuasive writing seeks to persuade the reader to take action or to bring about change. This type of writing involves a clear awareness of what arguments might most affect the audience being addressed. Writing persuasively also requires the use of such skills as analysis, inference, synthesis, and evaluation.

MCAS

The English language arts are organized into four Strands, or content areas: Language, Reading and Literature, Composition, and Media. The Composition strand consists of 7 learning standards:

- **Writing:** Students will write with a clear focus, coherent organization, and sufficient detail.
- **Consideration of Audience and Purpose:** Students will write for different audiences and purposes.
- **Revising:** Students will demonstrate improvement in organization, content, paragraph development, level of detail, style, tone, and word choice (diction) in their compositions after revising them.
- **Standard English Conventions:** Students will use knowledge of Standard English conventions in their writing, revising, and editing.
- **Organizing Ideas in Writing:** Students will organize ideas in writing in a way that makes sense for their purpose.
- **Research:** Students will organize ideas in writing in a way that makes sense for their purpose.
- **Evaluating Writing and Presentations:** Students will develop and use appropriate rhetorical, logical, and stylistic criteria for assessing final versions of their compositions or research projects before presenting them to varied audiences.

Test Construction

NAEP

- **Matrix sampling:** each student takes a subset of the total set of questions (2 of 17 tasks).

MCAS

- At each grade, every student gets the same writing prompt.

Type of Questions

NAEP

- The target percentage of assessment time is equally distributed among the three purposes for writing (Narrative 33%, 6 tasks; informative: 33%, 6 tasks; Persuasive, 33%, 5 tasks)

MCAS

- One writing prompt per grade:
Grade 4: personal narrative
Grade 7: expository essay
Grade 10: literacy analysis

Test Questions release

NAEP

- Only selected writing questions are released to public. For current year and historical released test questions, please visit:
<http://nces.ed.gov/nationsreportcard/itmrls/>

MCAS

- Writing prompts are released to public. For current year and historical released writing prompts, please visit:
<http://www.doe.mass.edu/mcas/testitems.html>

Testing Administration

2007 NAEP

Same for National NAEP, State NAEP, and Trial Urban District Assessment (TUDA) NAEP

Testing Date: 1/22/2007 – 3/2/2007

Testing Time: 50 minutes (two 25-minutes)

Test Grade: Grades 8 & 12 (National only)

Test Administering: The NAEP Representative from NAEP data collection contractor is responsible for all assessment activities including coordinating, conducting, and sending test materials to the scoring facility.

Test Sequence: sampled students received a specific assignment, such as position paper, story, or letter.

2007 MCAS

Testing Date: 3/20/2007 (make-up 4/2/2007)

Testing Time: Un-timed, with two sessions

Test Grade: Grades 4, 7, & 10

Test Administering: School teachers/personnel are responsible for all assessment activities.

Test Sequence: All students take the same test in the same classroom.

Scoring

NAEP

- Student responses are evaluated by expert scorers on a 6-level scoring criteria that described 6 performance ratings: unsatisfactory, insufficient, uneven, sufficient, skillful, and excellent.
- The elements of writing in scoring guides include content, purpose, organization, sentence structure, and conventions (grammar, spelling, punctuation)
- Student responses were scored as first drafts, not as polished pieces of writing.

MCAS

- Student compositions are independently scored by two scorers on the following criteria: (1) a score of 1–6 in idea/topic development, and (2) a score of 1-4 for the use of Standard English writing conventions. Students receive the sum of the scores from each of the two readers.
- The instructions urge the students to write a draft. However, some students may produce multiple drafts. Only the final version is scored.

Data Availability

NAEP

- No student-level results
- No school-level results
- No district-level results (except TUDA)
- Not designed to assess a specific curriculum

MCAS

- Student-level results
- School-level results
- District-level results
- Designed to measure the state's curriculum

Reporting

Performance Standard

NAEP

Three Achievement Levels:

- **Advanced:** Represents superior performance
- **Proficient:** Represents solid academic performance for each grade assessed
- **Basic:** Denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.

MCAS

Performance Level is no available. The writing prompt score is a part of overall ELA score.

Scale Score

NAEP

- **Range:** 0 – 300
- **Scale Score Corresponding to Performance Level:** varies by test grade

Performance Level:

	<u>Grade 8</u>
Advanced	224 – 300
Proficient	173 – 223
Basic	114 – 172
<i>Below Basic*</i>	0 – 113

* *Below Basic is not an Achievement level*

- Average scale scores cannot be compared across grades.
- The scale summarizes student performance across all three purposes for writing (narrative, informative and persuasive)

MCAS

No scale score is reported. Two raw scores are reported:

- Idea/topic development: 2 -12 points
- Standard English writing conventions: 2-8 points.

These raw scores are a part of overall ELA score.

Interpreting Results

NAEP

- The NAEP results as reported in average scores and percentages are **estimates** because they are based on samples rather than the entire population(s).
- Differences in scores must be statistically significant in order to report a difference between groups.

MCAS

- Comparisons of performance on writing scores across years must be made with caution because it is only based on one specific writing task.

Additional Information

NAEP

The Nation’s Report Card (NAEP) (NCES)
 National Center for Education Statistics
 U.S. Department of Education
 1990 K Street, NW
 Washington, DC 20006
 Phone: (202) 502-7300
 Web site:
<http://nces.ed.gov/nationsreportcard/>

MCAS

The Massachusetts Department of
 Elementary and Secondary Education
 Student Assessment Services
 350 Main Street
 Malden, MA 02148
 Phone: (781) 338-3616
 Web site: <http://www.doe.mass.edu/MCAS>

Appendix C

2007 NAEP Writing Assessment Sample of Released Prompts for Grade 8

Grade 8 – Narrative Prompt - “William Carlos Williams poem about missing plums”

Imagine that you have just come into your kitchen and that the poem below is a note left for you on the kitchen table. Who wrote the note? How do you feel? What do you do?

Write a story about what happens next.

THIS IS JUST TO SAY

I have eaten
The plums
That were in
the icebox*

and which
you were probably
saving
for breakfast

Forgive me
They were delicious
So sweet
And so cold

*"icebox" is another word for refrigerator.

“This is just to Say” by William Carlos Williams, from Collected Poems: 1909-1939, volume 1. Copyright © 1938 by New Directions Publishing Corp. Reprinted by permission of New Directions Publishing Corp.

Grade 8 – Informative Prompt - “Describe what a backpack is and how it functions”

Open the envelope labeled E that you have been given. Take out the letter from Rina and read it. Rina, who wrote the letter, is coming to a school in America for the first time and needs to know what a backpack is.

Write a letter back to Rina. In your letter, include a clear description of a backpack and explain in detail what she should keep in it. Remember, the more information Rina has, the better prepared she'll be to start eighth grade.

Grade 8 – Persuasive Prompt – “Letter to friend to convince them to try something new”

Imagine that one of your friends never wants to try anything new. Whether it is a question of what to eat, what to wear, what to do, what to read, or what to watch, your friend always chooses what is familiar – “the same old thing.”

Write a letter to your friend convincing him or her to try something new. Be sure to describe what your friend should try and explain why your friend would like it.

Appendix D

Average Scale Scores and Achievement Level in 2007 NAEP Grade 8 Writing by Selected Race/Ethnicity Categories and Jurisdiction

Table A-16. Average scores and achievement-level results in NAEP writing for eighth-grade public school students, by selected race/ethnicity categories and jurisdiction: 2007

Race/ethnicity and jurisdiction	Average scale score	Percentage of students			
		Below <i>Basic</i>	At or above <i>Basic</i>	At or above <i>Proficient</i>	At <i>Advanced</i>
White					
Nation (public)	162	8	92	39	2
Large central city	162	9	91	39	3
Atlanta	176	5	95	58	4
Austin	173*,**	5*	95*	53*,**	5
Boston	173*,**	6	94	52*,**	8
Charlotte	173*,**	3*,**	97*,**	52*,**	4
Chicago	170	8	92	54**	4
Cleveland	142*,**	14	86	13*,**	#
Houston	171*,**	4	96	46	4
Los Angeles	160	9	91	37	2
New York City	167	9	91	46	3
San Diego	167	7	93	47	3
Black					
Nation (public)	140*	20*	80*	15*	#
Large central city	138**	22**	78**	13**	#
Atlanta	142	17	83	16	#
Austin	130**	32**	68**	12	1
Boston	141	21	79	16	#
Charlotte	144*	17	83	17	#
Chicago	138	22	78	15	#
Cleveland	132*,**	25**	75**	7**	#
Houston	140	20	80	15	#
Los Angeles	129*,**	30	70	8**	#
New York City	140	21	79	15	#
San Diego	144	20	80	19	#
Hispanic					
Nation (public)	141*	21*	79*	17*	#*
Large central city	137**	24**	76**	14**	#**
Atlanta	‡	‡	‡	‡	‡
Austin	131**	30**	70**	12**	#
Boston	138	23	77	14	#
Charlotte	142	23	77	21	1
Chicago	148*,**	14*,**	86*,**	22*	#
Cleveland	133	28	72	10	#
Houston	138	22	78	13	#
Los Angeles	133*,**	25**	75**	9*,**	#
New York City	137	27**	73**	18	1
San Diego	129*,**	34*,**	66*,**	11**	#

See notes at end of table.

Table A-16. Average scores and achievement-level results in NAEP writing for eighth-grade public school students, by selected race/ethnicity categories and jurisdiction: 2007—Continued

Race/ethnicity and jurisdiction	Average scale score	Percentage of students			
		Below <i>Basic</i>	At or above <i>Basic</i>	At or above <i>Proficient</i>	At <i>Advanced</i>
Asian/Pacific Islander					
Nation (public)	166*	8	92	45*	5
Large central city	160**	12	88	40**	3
Atlanta	‡	‡	‡	‡	‡
Austin	‡	‡	‡	‡	‡
Boston	174	4	96	55	5
Charlotte	‡	‡	‡	‡	‡
Chicago	‡	‡	‡	‡	‡
Cleveland	‡	‡	‡	‡	‡
Houston	171	5	95	47	5
Los Angeles	160	7	93	35	2
New York City	167	10	90	49	4
San Diego	165	8	92	44	2

Rounds to zero.

‡ Reporting standards not met. Sample size is insufficient to permit a reliable estimate.

* Significantly different ($p < .05$) from large central city public schools in 2007.

** Significantly different ($p < .05$) from nation (public schools) in 2007.

NOTE: Black includes African American, Hispanic includes Latino, and Pacific Islander includes Native Hawaiian. Race categories exclude Hispanic origin. Results are not shown for students whose race/ethnicity was American Indian/Alaska Native or unclassified. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2007 Trial Urban District Writing Assessment.

Appendix E

Average Scale Scores and Achievement Level in 2007 NAEP Grade 8 Writing by Various Student Groups and Jurisdiction

By Gender:

Table A-17. Average scores and achievement-level results in NAEP writing for eighth-grade public school students, by gender and jurisdiction: 2007

Gender and jurisdiction	Average scale score	Percentage of students			
		Below <i>Basic</i>	At or above <i>Basic</i>	At or above <i>Proficient</i>	At <i>Advanced</i>
Male					
Nation (public)	144*	18*	82*	20*	1
Large central city	136**	26**	74**	14**	#
Atlanta	136**	24	76	12**	#
Austin	135**	29**	71**	18	1
Boston	138**	24**	76**	15	1
Charlotte	143*	20*	80*	18	#
Chicago	136**	25**	75**	14**	#
Cleveland	124*,**	34*,**	66*,**	4*,**	#
Houston	135**	27**	73**	12**	#
Los Angeles	129*,**	30**	70**	8*,**	#
New York City	136**	28**	72**	16	1
San Diego	137**	27**	73**	18	#
Female					
Nation (public)	164*	7*	93*	41*	3*
Large central city	155**	11**	89**	30**	2**
Atlanta	153**	10	90	26**	1
Austin	157**	14**	86**	35*,**	4
Boston	160*,**	10	90	35**	4
Charlotte	167*	5*	95*	43*	4
Chicago	157**	9	91	31**	1
Cleveland	143*,**	13**	87**	13*,**	#
Houston	150**	12**	88**	23*,**	1**
Los Angeles	145*,**	15**	85**	18*,**	1
New York City	156**	13**	87**	34**	2
San Diego	158**	14**	86**	38*	2

Rounds to zero.

* Significantly different ($p < .05$) from large central city public schools in 2007.

** Significantly different ($p < .05$) from nation (public schools) in 2007.

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2007 Trial Urban District Writing Assessment.

By Status as Students with Disabilities:

Table A-19. Average scores and achievement-level results in NAEP writing for eighth-grade public school students, by status as students with disabilities (SD) and jurisdiction: 2007

SD status and jurisdiction	Average scale score	Percentage of students			
		Below <i>Basic</i>	At or above <i>Basic</i>	At or above <i>Proficient</i>	At <i>Advanced</i>
SD					
Nation (public)	118*	46*	54*	6*	#
Large central city	112**	54**	46**	4**	#
Atlanta	105	60	40	3	#
Austin	111	52	48	6	#
Boston	121*	41	59	4	#
Charlotte	120*	46	54	6	#
Chicago	107**	58**	42**	5	#
Cleveland	96*,**	73*,**	27*,**	1	#
Houston	110	56	44	3	#
Los Angeles	105**	61**	39**	2**	#
New York City	112	52	48	1	#
San Diego	108	59	41	5	#
Not SD					
Nation (public)	159*	9*	91*	33*	2*
Large central city	149**	15**	85**	24**	1**
Atlanta	148**	13	87	21**	#
Austin	151**	17**	83**	29*,**	2
Boston	154*,**	13**	87**	29*,**	3
Charlotte	159*	8*	92*	34*	2
Chicago	153**	10*	90*	26**	1
Cleveland	138*,**	17**	83**	10*,**	#
Houston	145**	16**	84**	19*,**	1**
Los Angeles	140*,**	19*,**	81*,**	14*,**	#
New York City	152**	15**	85**	29*,**	2
San Diego	151**	17**	83**	29*,**	1

Rounds to zero.

* Significantly different ($p < .05$) from large central city public schools in 2007.

** Significantly different ($p < .05$) from nation (public schools) in 2007.

NOTE: The results for students with disabilities are based on students who were assessed and cannot be generalized to the total population of such students. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2007 Trial Urban District Writing Assessment.

By Status as English Language Learners:

Table A-20. Average scores and achievement-level results in NAEP writing for eighth-grade public school students, by status as English language learners (ELL) and jurisdiction: 2007

ELL status and jurisdiction	Average scale score	Percentage of students			
		Below <i>Basic</i>	At or above <i>Basic</i>	At or above <i>Proficient</i>	At <i>Advanced</i>
ELL					
Nation (public)	120*	42*	58*	5*	#
Large central city	112**	51**	49**	3**	#
Atlanta	‡	‡	‡	‡	‡
Austin	100*,**	65*,**	35*,**	1	#
Boston	102*,**	68*,**	32*,**	1	#
Charlotte	126*	38	62	6	#
Chicago	117	45	55	4	#
Cleveland	‡	‡	‡	‡	‡
Houston	102*,**	65*,**	35*,**	1	#
Los Angeles	113**	48**	52**	2**	#
New York City	101*,**	67*,**	33*,**	2	#
San Diego	107**	59**	41**	1**	#
Not ELL					
Nation (public)	156*	11*	89*	32*	2*
Large central city	149**	15**	85**	24**	1**
Atlanta	145**	16	84	19**	#
Austin	152**	16**	84**	30*	2
Boston	154*	12	88	28**	2
Charlotte	157*	10*	90*	33*	2
Chicago	148**	16**	84**	23**	1
Cleveland	134*,**	22*,**	78*,**	9*,**	#
Houston	147**	14**	86**	19**	1**
Los Angeles	146**	13	87	18*,**	#
New York City	151**	15	85	28	2
San Diego	157*	11*	89*	33*	2

Rounds to zero.

‡ Reporting standards not met. Sample size is insufficient to permit a reliable estimate.

* Significantly different ($p < .05$) from large central city public schools in 2007.

** Significantly different ($p < .05$) from nation (public schools) in 2007.

NOTE: The results for English language learners are based on students who were assessed and cannot be generalized to the total population of such students. Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2007 Trial Urban District Writing Assessment.

By Eligibility for Free/Reduced-Price School Lunch:

Table A-18. Average scores and achievement-level results in NAEP writing for eighth-grade public school students, by eligibility for free/reduced-price school lunch and jurisdiction: 2007

Eligibility status and jurisdiction	Average scale score	Percentage of students			
		Below <i>Basic</i>	At or above <i>Basic</i>	At or above <i>Proficient</i>	At <i>Advanced</i>
Eligible					
Nation (public)	141*	20*	80*	17*	#
Large central city	138**	23**	77**	15**	#
Atlanta	140	20	80	14	#
Austin	128*,**	33*,**	67*,**	9*,**	#
Boston	144*	18	82	18	1
Charlotte	141	20	80	15	#
Chicago	142	19*	81*	18	#
Cleveland	133*,**	23	77	9*,**	#
Houston	137	22	78	13	#
Los Angeles	133*,**	25**	75**	9*,**	#
New York City	144*	21	79	22*,**	1
San Diego	133**	31*,**	69*,**	14	#
Not eligible					
Nation (public)	164*	7*	93*	40*	3
Large central city	159**	11**	89**	36**	2
Atlanta	162	7	93	38	2
Austin	168*,**	7	93	47*	4
Boston	161	15**	85**	41	6
Charlotte	169*	5*	95*	46*	4
Chicago	169*	8	92	50*	3
Cleveland	‡	‡	‡	‡	‡
Houston	159	10	90	35	2
Los Angeles	150**	15	85	26	1
New York City	167	8	92	45	5
San Diego	163	9	91	42	3
Information not available					
Nation (public)	149	15	85	25	2
Large central city	147	16	84	23	1
Atlanta	‡	‡	‡	‡	‡
Austin	‡	‡	‡	‡	‡
Boston	‡	‡	‡	‡	‡
Charlotte	‡	‡	‡	‡	‡
Chicago	‡	‡	‡	‡	‡
Cleveland	‡	‡	‡	‡	‡
Houston	‡	‡	‡	‡	‡
Los Angeles	147	16	84	23	1
New York City	‡	‡	‡	‡	‡
San Diego	‡	‡	‡	‡	‡

Rounds to zero.

‡ Reporting standards not met. Sample size is insufficient to permit a reliable estimate.

* Significantly different ($p < .05$) from large central city public schools in 2007.

** Significantly different ($p < .05$) from nation (public schools) in 2007.

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2007 Trial Urban District Writing Assessment.