Education Policy Analysis Archives

Volume 9 Number 42 **October 16, 2001** **ISSN 1068-2341**

A peer-reviewed scholarly journal
Editor: Gene V Glass, College of Education
Arizona State University

Articles appearing in **EPAA** are abstracted in the *Current Index to Journals in Education* by the ERIC Clearinghouse on Assessment and Evaluation and are permanently archived in *Resources in Education*.

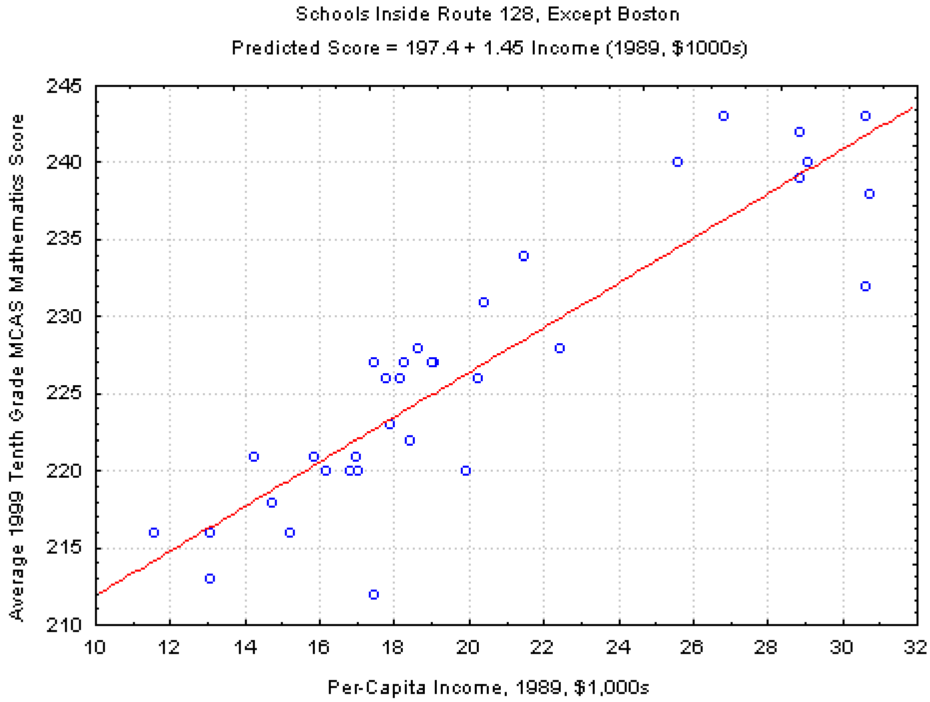
Significance of Test-based Ratings for Metropolitan Boston Schools

Craig Bolon
Planwright Systems Corporation (USA)

Citation: Bolon, C. (2001, October 16). Significance of Test-based Ratings for Metropolitan Boston Schools. *Education Policy Analysis Archives*, *9*(42). Retrieved [date] from http://epaa.asu.edu/epaa/v9n42/.

Abstract

In 1998 Massachusetts began state-sponsored, annual achievement testing of all students in three public school grades. It has created a school and district rating system for which scores on these tests are the sole factor. It proposes to use tenth-grade test scores as a sole criterion for high school graduation, beginning with the class of 2003. The state is treating scores and ratings as though they were precise educational measures of high significance. A review of tenth-grade mathematics test scores from academic high schools in metropolitan Boston showed that statistically they are not. Community income is strongly correlated with test scores and accounted for more than 80 percent of the variance in average scores for a sample of Boston-area communities:



Once community income was included in models, other factors--including percentages of students in disadvantaged populations, (Note 1) percentages receiving special education, percentages eligible for free or reduced price lunch, percentages with limited English proficiency, school sizes, school spending levels, and property values--all failed to associate substantial additional variance. Large uncertainties in residuals of school-averaged scores, after subtracting predictions based on community income, tend to make the scores ineffective for rating performance of schools. Large uncertainties in year-to-year score changes tend to make the score changes ineffective for measuring performance trends.

Contents

Section 1: Background

A. State Testing in Massachusetts Public Schools

The most recent form of state testing in Massachusetts public schools is the Massachusetts Comprehensive Assessment System (MCAS), a set of achievement tests sponsored and produced by the Massachusetts Department of Education and administered in the spring of years beginning in 1998 (see Appendix 1 and Bolon, 2000). These tests have four sections: English language arts, mathematics, science and technology, and history and social science. For the years 1998-2000, tests were administered in grades 4, 8 and 10. Beginning in April, 2001, the former grade 4 test sections were divided between grades 4 and 5; new tests have been added in grades 3, 6 and 7.

MCAS tests are loosely timed and include questions in multiple choice, short answer and extended answer formats; they are provided in English and Spanish. All public school students are required to take MCAS tests; there are no "opt-out" provisions. Students taught in parochial schools, other private schools, home schooling and out-of-state schools are not required to take or pass MCAS. For 1998 through 2000, the Department of Education has published test questions used for scoring approximately six months after test administration (see Mass. DoE, 2000g, for example). It has produced new test forms each year. According to current plans, starting with the class of 2003 minimum scores on the English language arts and mathematics sections will be required to graduate from high school (Mass. DoE, 1999f) and to enroll at state colleges, except for MCAS test preparation courses at two-year colleges.

The Massachusetts Department of Education publishes MCAS results as scaled scores in a range of 81 scale points, using the integers 200 through 280 (Mass. DoE, 2000h). The Department assigns labels it calls "performance levels" to four scaled score intervals (Mass. DoE, 1998b) and currently considers 220 the minimum passing scaled score on all test sections (Mass. DoE, 1999f). The Department has not fully disclosed details of assigning scale factors, assuring consistent scores across test forms or assuring that scores quantitatively reflect published academic standards. It has not published distributions of either raw scores or scaled scores. It has released limited information about test design and properties in "technical reports" for the years 1998 (Mass. DoE, 1999c) and 1999 (Mass. DoE, 2000i). After an independent analysis of score averages by "racial" and "ethnic" categories for 1998 (Uriarte and Chavez, 1999), the Department published its own analysis of this type for 1999 (Mass. DoE, 2000c).

Massachusetts tests appear to rank near the high end of state achievement tests in difficulty, although failure rates are lower than those for some tests used in Arizona and Virginia. As in several other states, substantially higher failure rates are found on mathematics than on language tests in high schools. Since the tenth-grade version of the mathematics test section sets the graduation threshold for most students, its scores have been used as subjects for these studies.

B. Schools in the Boston Metropolitan Area

Metropolitan Boston is diverse. Besides the City of Boston it includes many smaller municipalities, all operating their own school systems. These studies consider communities inside Route 128, a highway designed in the late 1940s (now an Interstate), enclosing areas within about 9-12 miles from Boston's government center--that is, Boston and its inner and middle suburbs. They share a public transit system, several public and private utilities, and an economy dominated by service industries. They include poverty areas, concentrations of wealth, middle-income communities, prosperous suburban towns, a few medium-sized cities and one large city. The areas are bounded by the Massachusetts Bay and Atlantic Ocean to the east, Salem and Peabody to the north, Waltham and Newton to the west, and Braintree and Quincy to the south (see Metropolitan Area, 1997).

Schools in the Boston metropolitan area are also diverse. These studies, focusing on testing for graduation, consider only high schools. While a majority of the area's population of high-school age attends public schools, (Note 2) a substantial proportion attends parochial schools that began to be established by the Roman Catholic Church more than 150 years ago. A smaller fraction is taught in other private schools or though home schooling. The Metropolitan Council for Educational Opportunities (METCO), founded in 1963, uses state funding to help send over 3,000 Boston minority students to suburban schools (Orfield, et al., 1997).

Within public school systems there is also substantial diversity. All communities must support regional "vocational," "technical" and "agricultural" high schools. Some such schools began as "manual training" schools in the 1800s. Some communities have closed their local vocational schools; some have merged them with their academic schools. These studies look in detail only at academic schools, because the curriculum of vocational schools is substantially different and is not designed to prepare students for MCAS tests, an issue of controversy (Nicodemus, 2000). For purposes of these studies there are difficulties with a few communities, including Cambridge, Quincy, Revere and Waltham, which provide vocational education in the same facilities as academic programs (Mass. DoE, 2000f). I chose to include such schools in these studies while noting their special characteristics.

Several communities also operate experimental schools, including "pilot schools" in Boston and "charter schools" in several communities (Partee, 1997, and Wood, 1999), as regulated under the Massachusetts Education Reform Act of 1993. All that offer ninth-grade curriculum and above are smaller than the regular academic schools. These schools provide motivational environments and may exercise indirect forms of student selection that differentiate them from other public schools. Primarily because of concerns about small sample sizes, schools with fewer than 100 students per grade are excluded from these studies. So far no experimental school is that large.

The City of Boston presents a unique situation. Of its large academic high schools, three are exam schools: the Boston Latin School (founded in 1635) and the more recent Latin Academy (formerly Girls Latin) and O'Bryant School of Mathematics and Science (formerly Boston Technical). These draw away many Boston students who tend to score well on achievement tests, promoting a longstanding social stratification in Boston schools. Over half the students at Boston Latin come to it from parochial and other private schools (Daley, 1997); some say those students would not otherwise attend Boston schools. However, other public school students who are not admitted leave the district for high school. Starting in 1975, because of federal court orders to desegregate, exam school admission policies included a 35 percent set-aside for African American and Latino students, maintained voluntarily after 1987. As a result of another federal court decision (McLaughlin, 1996), this approach was weakened in 1997. As with academic schools that provide vocational education, the Boston exam schools are included in these studies, but their special characteristics are noted.

C. Statewide MCAS Test Results

Table 1-1 shows that statewide, tenth-grade MCAS test scores have remained nearly constant in English language arts and in science and technology for the years 1998-2000, while scores in mathematics have risen substantially (Mass. DoE, 2000h). (Tenth-grade tests were not given in history and social science.)

Table 1-1
MCAS Statewide Results, 1998-2000

Section	Year	Average	% Level 4	% Level 3	% Level 2	% Level 1
English	2000	229	7	29	30	34
English	1999	229	4	30	34	32
English	1998	230	5	33	34	28
Math.	2000	228	15	18	22	45
Math.	1999	222	9	15	23	53
Math.	1998	222	7	17	24	52
Science	2000	226	3	23	37	37
Science	1999	226	3	21	39	38
Science	1998	225	1	21	42	36

Source of data: Mass. DoE, 2000h

Table 1-1 reflects Massachusetts Department of Education practice of recording students absent for a test section as scoring 200 and in Level 1, the lowest level (Mass. DoE, 2000h, Table 11 footnote). An undisclosed fraction of students were excluded from testing because of special conditions and are not counted in this report; others may have been provided with an alternative assessment. As currently planned, students in Level 1 will be ineligible to graduate from high school as of 2003. Based on this record of scores, about half of all Massachusetts public school students are at risk of being denied graduation.

The labels of the four "performance levels" designated by the 1998 Board of Education(Note 3) for reporting MCAS results are:

- Level 4, "**Advanced**"
- Level 3, "**Proficient**"
- Level 2, "**Needs Improvement**"
- Level 1, "**Failing**"

Although these levels have qualitative descriptions (Mass. DoE, 1998b), there are no quantitative links to levels of achievement specified in academic standards; content of standards has not been prioritized; nor have standards been promulgated through state regulations, as anticipated by law. (Note 4) Although Massachusetts law requires "competency determination" in mathematics, science and technology, history and social science, foreign languages and English, (Note 5) Massachusetts laws and regulations continue to require only US history and physical education as subjects of instruction. Massachusetts tries to set legal standards for learning indirectly (Note 6) through MCAS tests, procedures to set scale factors, and regulations for minimum scaled scores. It lacks corresponding legal commitments for instruction. It has made major changes to "curriculum frameworks" every few years (Mass. DoE, 2000k) and has not provided reasonable spans of time for instruction to catch up before using revised "curriculum frameworks" as a basis for revised MCAS tests. Its teachers, parents and students cannot find out exactly what must be learned in order to meet minimum standards for high school graduation. The 1993 Education Reform Act left several such problems; few have been addressed yet by the Massachusetts legislature or Board of Education.

Students with disabilities (also called special education students) and students with limited English proficiency (LEP students) tend to receive drastically lower MCAS scores than other students, although some students with disabilities are soon to be provided alternate assessments (Mass. DoE, 2000l), and some LEP students have been able to take tests in Spanish (Mass. DoE, 2000d). The Department of Education has not disclosed the fractions of students who are eligible for or have utilized its special accommodations, although it has published statewide summary data using these student categories (Mass. DoE, 2000i, Table 14.5). Most minority students also receive lower scores than other students. The Department of Education has published 1999 statewide and district summary data for students categorized as "African American / Black," "Asian or Pacific Islander," "Hispanic / Latino," "Native American," "White" and "Mixed" (Mass. DoE, 2000c, Tables 5-10). As previously noted, most students in vocational programs receive lower MCAS scores than students in academic programs; this can readily be shown for the state's more than 30 vocational, technical and agricultural high schools (Appendix 2).

Based on the sources of information cited, Table 1-2 shows statewide impacts of these known risk factors on average 1999 tenth-grade mathematics scores and rates of failure.

Table 1-2
MCAS 1999 Grade 10 Math Scores by Risk Factors

Category	Average Score	Percent Failing
All students	222	53
Students with disabilities	203	92
Limited English proficiency	206	84
African American	209	80
Hispanic / Latino	208	85
Native American	211	77

Vocational, technical, agricultural	210	78

Source of data: Mass. DoE, 2000i

The Department of Education has not reported scores classified by other potential risk factors on which it collects information. These include: Gender of students, Tests taken in Spanish or as alternate assessments, Free or reduced price lunches, as indicators of poverty, Schools with large class sizes, especially in early grades, Students retained below grade or placed below grade level, Teachers who lack certification in their subjects of instruction.

There is also little published information about combinations of risk factors. However, since the Department of Education lists regional vocational, technical and agricultural schools as separate districts in its reports of MCAS results, it is possible to use their categories of minority students (Appendix 2). For those schools for which categories are reported, results are shown in Table 1-3.

Table 1-3
MCAS 1999 Grade 10 Math Scores by Combined Risk Factors

Combined Category	Average Score	Percent Failing
Vocational + African American	203	97
Vocational + Hispanic / Latino	205	95

Sources of data: Mass. DoE, 2000i, Mass. DoE, 2000h

While the results in Table 1-3 are not strictly comparable with Table 1-2, because not all the schools and categories can be found in published data, they indicate that factors can combine to worsen the scores of students with more than one risk factor.

D. Test Score Studies

Recent studies question assumptions that "high-stakes" tests like MCAS can provide valid measures of either student achievement or school performance, showing gains on them that are not matched by gains on other tests for closely related educational content (Haney, 2000, and Klein, et al., 2000). Political environments of "high-stakes" tests create heavy pressure to improve scores, regardless of underlying educational progress. For "low-stakes" tests aimed at measuring long-term trends, like those of the federal NAEP, it has been shown that "family variables explain most of the variance across scores in states" (Grissmer, et al., 2000, Chapter 9). Individual and longitudinal studies demonstrate strong influences of parenting practices, family structure, parent education and degrees of poverty on cognitive development (for example, Smith, et al., 1997). Other longitudinal and cross-sectional studies show cumulative responses of test scores to educational environments (for example, Phillips, et al., 1998, and Ferguson, 1998). However, the data generally available for test score research fail to capture much of the critical information needed to understand development of cognitive abilities and educational achievement in the settings of public schools.

MCAS test scores have already been the subject of several attempts to explain, predict or interpret them (Mass. DoE, 2001, Gaudet, 2001, Tuerck, 2001a, and Tuerck, 2001b). These prior MCAS test score studies fall into three main categories: 1) Trends studies of year-to-year and multi-year changes; 2) Effects studies involving social factors for the population; 3) Effects studies involving operating factors for the schools.

Research on scores from school-based standard tests suggests that many such studies are likely to yield results of low significance. Grissmer, et al., 2000, among others, show that:

- Real year-to-year changes in average student performance, as assessed by conventional tests, are relatively small; they can easily be masked by statistical uncertainties.
- Social factors are strongly associated with test scores.
- Self-reported social information tends to have high error and omission rates.
- Census and other community-based social information often includes confounding factors that require adjustment to reflect the households for a school population.
- Uncategorized school spending is only weakly associated with test scores.

The MCAS test score studies cited use scores and statistical data to estimate the performance of schools or districts according to simple formulas, unsupported by other evidence. They frequently present results in a table that is ranked or can be ranked like the teams in a sports league. The "league table" approach to presenting such results begs the question of whether the ordering of schools or districts and the differences in performance estimates have educational significance, that is, whether such rankings may instead be largely matters of chance or be associations with factors other than school performance. This article presents a trends study and an effects study I conducted to explore the significance that can be associated with such results.

E. Sources of Data

The school characteristics used in these studies are taken from information reported by public schools to the Massachusetts Department of Education for 1999 and published by the Department (Mass. DoE, 2000f). MCAS test scores summarized by schools are from 1998-2000 Department reports (Mass. DoE, 2000h). Other information is published by the Department for school districts, including program budgets and percentages of special education students. Information for census tracts and communities is available from the US Bureau of the Census and other sources. Data analysis for these studies focuses on information associated with individual schools because aggregate information for school districts or general populations can mask school characteristics. Data used in these studies are reproduced in Appendix 3 and Appendix 4; interested readers can confirm them at the sources and can repeat these studies or perform other analysis with them.

The Department of Education and the school districts collect other potentially useful information that is not currently published. Of particular interest are data on class size and teacher preparation. Recent research has shown significant association of educational achievement as measured by "low-stakes" tests with small class size in elementary schools (Nye, et al., 1999, and Krueger, 1999) and with teacher certification and education (Darling-Hammond, 2000), after adjustments for student backgrounds. Studies of the development of cognitive abilities cast doubt on whether other information currently published by government sources about population and economic characteristics in large geographical areas would substantially improve the understanding of test scores.

A. Trends Study of Variability

This study considers 47 academic high schools in 32 metropolitan Boston communities through the average tenth-grade MCAS mathematics test scores recorded for years 1998-2000. Achievement tests in mathematics typically require substantial skill at language interpretation (see, for example, Gipps and Murphy, 1994, Chapter 6, p. 183). Haney, 2000, in a study of another state, found stronger correlations of state mathematics test scores with grades in English than with grades in math. As previously noted, the tenth-grade mathematics test is used in this study of significance because it sets a graduation threshold for most students.

Test boycotts have been organized by students in several schools each year (Steinberg, 2000), involving 10 to 31 percent of students in 19 cases out of the 141 test samples. To be able to compare average scores of schools more accurately, the average scores reported by the Department of Education have been adjusted by removing the scores of 200 that were assigned to students who did not take the test, averaging only scores of students who participated.

Table 2-1 shows changes in schools' average scores (Appendix 3) between 1998 and 1999 and between 1999 and 2000, expressed in units of scale points and of standard deviations.

Table 2-1
MCAS Grade 10 Math Test Score Changes by School, 1998-2000

		Changes 1998-1999		Changes 1999-2000	
City or Town	**High School**	**Points**	**Delta**	**Points**	**Delta**
Arlington	Arlington	4	3	5	-1
Belmont	Belmont	-2	-4	4	-2
Boston	Boston High	1	0	6	0
Boston	Brighton	2	1	3	-4
Boston	Charlestown	-1	-2	4	-2
Boston	Dorchester	-2	-3	1	-5
Boston	East Boston	1	0	5	-1
Boston	Hyde Park	0	-1	1	-5
Boston	Jeremiah Burke	4	3	3	-3
Boston	South Boston	0	-1	5	-1
Boston	The English High	1	0	4	-2
Boston	West Roxbury	3	2	5	-1
Boston Exam	Boston Latin	8	10	8	3
Boston Exam	Latin Academy	3	2	19	17
Boston Exam	O'Bryant Science	4	4	8	3
Braintree	Braintree	-1	-3	13	10
Brookline	Brookline	2	1	5	-1
Cambridge	Rindge & Latin*	-2	-5	-1	-10

Table 2-5
Factor Correlations for MCAS Grade 10 Math Test Scores

Factor	**A**	**B**	**C**	**D**	**E**	**F**	**G**
A. School population, average per grade	1.00	-.11	.24	.02	-.03	-.01	.16
B. Percent African American	-.11	1.00	.07	.56	.80	.80	.28
C. Percent Asian or Pacific Islander	.24	.07	1.00	.08	.13	.23	.27
D. Percent Hispanic / Latino	.02	.56	.08	1.00	.77	.84	.65
E. Percent limited English proficiency	-.03	.80	.13	.77	1.00	.88	.50
F. Percent free or reduced price lunch	-.01	.80	.23	.84	.88	1.00	.60
G. Percent reduction, grades 9+10 to 11+12	.16	.28	.27	.65	.50	.60	1.00

Sources of data: Appendix 4, Statistica model

Some of the correlations in Table 2-5 are strong enough that multiple regression coefficients are likely to be unstable. Therefore a model was developed in stages, examining factors for significance.

The full model from the factors in Table 2-4 was first evaluated with weights proportional to numbers of test participants. It yielded two strong factors with low correlation (**C** and **E**): "Percent Asian or Pacific Islander" at $p<.02$, with a positive coefficient, and "Percent limited English proficiency" at $p<.002$, with a negative coefficient: Factors for school population and percent grade reduction had particularly small coefficients and low significance. They were removed, and a model with the remaining five factors then associated 67 percent of the variance and produced the factor weights shown in Table 2-6.

Table 2-6
5-Factor Model for 1999 MCAS Grade 10 Math Test Scores

Factor	**Coefficient**	**Standard Error**
Intercept, for all factors zero	229.4	1.936
B. Percent African American	0.047	0.104
C. Percent Asian or Pacific Islander	0.347	0.154
D. Percent Hispanic / Latino	-0.002	0.183
E. Percent limited English proficiency	-0.637	0.217
F. Percent free or reduced price lunch	-0.174	0.157

Sources of data: Appendix 3, Appendix 4, Statistica model

With model factors in Table 2-6, high factor weight and significance found in other studies for percentages of African American or Latino students disappear. Both factors have small coefficients and low significance. Statistical weight that might have been attached to these factors instead follows cultural and economic factors: "Percent limited English proficiency" and "Percent free or reduced price lunch." As an experiment, the model was rerun with the latter factors removed; only 57 percent of the variance was associated, and factor weights became those shown in Table 2-7.

Table 2-7
Racial and Ethnic Model for 1999 MCAS Grade 10 Math Test Scores

Factor	Coefficient	Standard Error
Intercept, for all factors zero	230.0	2.107
B. Percent African American*	-0.221	0.068
C. Percent Asian or Pacific Islander	0.219	0.156
D. Percent Hispanic / Latino*	-0.435	0.114

Sources of data: Appendix 3, Appendix 4, Statistica model

In Table 2-7, two "racial" or "ethnic" factors (marked *) have become significant at a $p<.05$ level. The coefficient for "Percent African American" has turned from positive to negative, and the coefficient for "Percent Hispanic / Latino" has become strongly negative. It seems likely that these two factors are acting as proxies for cultural and economic factors with more predictive power.

Residuals from the five-factor model of Table 2-6 are shown in Table 2-8. This include standard error estimates based on results from the trends study of Section 2A.

Table 2-8
Residuals for MCAS Grade 10 Math Test Scores, 5-Factor Model

City or Town	High School	Residual	Std. Error	Ratio
Arlington	Arlington	4.6	2.6	1.8
Belmont	Belmont	12.2	2.7	4.5
Boston	Boston High	-8.9	5.1	-1.7
Boston	Brighton	1.7	3.3	0.5
Boston	Charlestown	1.0	4.8	0.2
Boston	Dorchester	-1.6	4.3	-0.4
Boston	East Boston	2.1	4.2	0.5
Boston	Hyde Park	-6.1	4.8	-1.3
Boston	Jeremiah Burke	4.1	5.0	0.8
Boston	South Boston	-8.6	3.7	-2.3
Boston	The English High	11.4	5.0	2.3
Boston	West Roxbury	1.2	3.9	0.3
Boston Exam	Boston Latin	21.3	3.5	6.0
Boston Exam	Latin Academy	4.0	4.1	1.0
Boston Exam	O'Bryant Science	4.65	4.3	1.1
Braintree	Braintree	-1.78	2.5	-0.7
Brookline	Brookline	9.5	2.3	4.1
Cambridge	Rindge & Latin*	-6.3	4.0	-1.6

Chelsea	Chelsea	2.2	7.0	0.3
Dedham	Dedham	-1.6	3.1	-0.5
Everett	Everett*	-2.8	2.6	-1.1
Lexington	Lexington	4.4	2.8	1.6
Lynn	Classical	-6.5	3.3	-2.0
Lynn	English	-5.0	3.9	-1.3
Malden	Malden	-6.6	3.2	-2.1
Marblehead	Marblehead	3.4	3.2	1.1
Medford	Medford*	-7.4	2.7	-2.7
Melrose	Melrose	-2.7	2.7	-1.0
Milton	Milton	-1.7	3.1	-0.6
Newton	North*	8.6	2.3	3.7
Newton	South	10.0	2.8	3.6
Peabody	Veterans*	-7.2	2.5	-2.9
Quincy	North Quincy	-9.2	3.6	-2.5
Quincy	Quincy*	-14.1	3.1	-4.5
Revere	Revere*	-10.3	2.7	-3.9
Salem	Salem*	-1.5	3.1	-0.5
Saugus	Saugus	-2.8	2.9	-1.0
Somerville	Somerville*	2.7	4.7	0.6
Stoneham	Stoneham	-1.7	3.1	-0.5
Swampscott	Swampscott	11.8	3.2	3.7
Wakefield	Memorial	-2.0	2.8	-0.7
Waltham	Waltham*	-7.4	2.6	-2.8
Watertown	Watertown	4.7	3.1	1.5
Weymouth	Weymouth*	-7.0	2.3	-3.0
Winchester	Winchester	12.5	2.9	4.4
Winthrop	Winthrop*	-5.5	3.5	-1.6
Woburn	Woburn	-1.9	2.6	-0.7

* school providing vocational education

Sources of data: Appendix 3, Appendix 4, Statistica model

At first glance, some residuals in Table 2-8 look substantial, several scale points of difference from the average scores predicted by the model. However, residual ratios for most schools are within +/- 2 standard errors, not significant at a $p<.05$ level. Someone familiar with metropolitan Boston will recognize that schools with high and low residual ratios tend to be in high-income and low-income communities, respectively. It therefore seems likely that adding a factor for incomes can increase the predictive power of the model.

The most recent community income data were from the US Census of 1990, for 1989 per-capita income. Comparable 1999 income statistics were not yet available. The Massachusetts Department of Revenue could produce current community income statistics but has not done so; the state continues to use 1989 federal census data on incomes to apportion aid to public schools. After adding 1989 per-capita community income in $1,000s as a factor (Mass. DoR, 1999), without any attempt to adjust incomes so as to reflect school districts or student households, the model associates 80 percent of the statistical variance, and factor weights became those shown in Table 2-9.

Table 2-9
6-Factor Model for 1999 MCAS Grade 10 Math Test Scores

Factor	Coefficient	Standard Error
Intercept, for all factors zero	202.7	5.400
B. Percent African American	-0.020	0.083
C. Percent Asian or Pacific Islander*	0.371	0.121
D. Percent Hispanic / Latino	0.044	0.144
E. Percent limited English proficiency*	-0.695	0.171
F. Percent free or reduced price lunch	0.050	0.131
H. Per-capita community income (1989)*	1.186	0.230

Sources of data: Appendix 3, Appendix 4, Statistica model

Three factors in Table 2-9 (marked *) have substantial significance, at a $p<.005$ level or better, and three have very low significance. Factor weight has shifted from "Percent free or reduced price lunch" to "Per-capita community income (1989)," while "Percent limited English proficiency" retains a large coefficient and high significance. Dropping low-significance factors, the resulting three-factor model is shown in Table 2-10.

Table 2-10
3-Factor Model for 1999 MCAS Grade 10 Math Test Scores

Factor	Coefficient	Standard Error
Intercept, for all factors zero	204.9	4.446
C. Percent Asian or Pacific Islander	0.381	0.109
E. Percent limited English proficiency	-0.626	0.081
H. Per-capita community income (1989)	1.104	0.197

Sources of data: Appendix 3, Appendix 4, Statistica model

The three-factor model of Table 2-10 also associates 80 percent of the statistical variance. All of its factors are statistically significant at a $p<.001$ level.

For each school included in these studies, Table 2-11 presents adjusted average 1999 tenth-grade MCAS mathematics test scores and residuals from the three-factor statistical model of Table 2-10, with the uncertainties in average scores and residuals expressed as standard errors, based on the variance estimate calculated in the trends study of Section 2A.

Table 2-11
Residuals for MCAS Grade 10 Math Test Scores, 3-Factor Model

City or Town	High School	Average	Std. Error	Residual	Std. Error
Arlington	Arlington	234	2.1	4.2	2.4
Belmont	Belmont	243	2.2	6.7	2.7
Boston	Boston High	204	2.9	-10.3	3.1
Boston	Brighton	205	2.2	-0.1	2.9
Boston	Charlestown	206	2.7	-1.6	3.7
Boston	Dorchester	204	3.0	-0.5	3.5
Boston	East Boston	205	2.3	-1.0	2.9
Boston	Hyde Park	203	3.3	-5.0	3.7
Boston	Jeremiah Burke	208	2.7	5.1	3.4
Boston	South Boston	205	2.6	-7.8	3.1
Boston	The English High	204	2.3	9.7	3.7
Boston	West Roxbury	205	2.0	0.0	2.7
Boston Exam	Boston Latin	254	1.7	23.6	2.7
Boston Exam	Latin Academy	233	2.1	5.0	2.8
Boston Exam	O'Bryant Science	227	2.1	4.3	3.4
Braintree	Braintree	228	1.9	1.7	2.3
Brookline	Brookline	240	1.6	-0.5	2.7
Cambridge	Rindge & Latin*	220	1.6	-4.9	1.8
Chelsea	Chelsea	216	2.3	3.7	2.8
Dedham	Dedham	227	2.5	1.3	2.8
Everett	Everett*	221	1.9	2.3	2.5
Lexington	Lexington	238	1.7	-5.8	3.0
Lynn	Classical	216	2.0	-3.3	2.8
Lynn	English	213	2.1	0.9	2.5
Malden	Malden	221	2.0	-2.8	2.6
Marblehead	Marblehead	232	2.7	-6.3	3.6
Medford	Medford*	221	2.2	-2.3	2.5
Melrose	Melrose	226	2.1	-1.2	2.5
Milton	Milton	228	2.3	-2.0	2.6
Newton	North*	239	1.5	0.8	2.5

Newton	South	242	2.0	2.2	2.8
Peabody	Veterans*	220	1.8	-2.9	2.3
Quincy	North Quincy	227	1.9	-7.2	3.0
Quincy	Quincy*	212	2.3	-11.9	2.6
Revere	Revere*	218	1.9	-6.1	2.4
Salem	Salem*	220	2.1	0.7	2.5
Saugus	Saugus	226	2.2	0.8	2.6
Somerville	Somerville*	216	1.8	1.7	2.1
Stoneham	Stoneham	227	2.5	1.5	2.9
Swampscott	Swampscott	240	2.6	7.6	3.0
Wakefield	Memorial	227	2.2	0.6	2.6
Waltham	Waltham*	220	1.8	-4.0	2.1
Watertown	Watertown	231	2.6	4.3	2.8
Weymouth	Weymouth*	222	1.5	-3.8	2.1
Winchester	Winchester	243	2.3	2.8	3.2
Winthrop	Winthrop*	223	3.0	-1.4	3.3
Woburn	Woburn	226	2.0	1.3	2.4

* school providing vocational education
Sources of data: Appendix 3, Appendix 4, Statistica model

Stepwise analysis shows that combinations of the factors in the three-factor model of Table 2-10 associate statistical variance in the amounts listed in Table 2-12.

Table 2-12
Factor Combinations for 1999 MCAS
Grade 10 Math Test Scores

Factors	R^2
C	.01
E	.62
H	.47
C E	.65
C H	.52
E H	.74
C E H	.80

C. Percent Asian or Pacific Islander
E. Percent limited English proficiency
H. Per-capita community income (1989)

Sources of data: Appendix 3, Appendix 4, Statistica model

Table 2-12 shows that the major factors are "Percent limited English proficiency" and "Per-capita community income (1989)." Although statistically significant, "Percent Asian or Pacific Islander" is a weak cofactor, associating only 1 percent of the variance by itself.

The three-factor model of Table 2-10 was evaluated for predictions of 1998 and 2000 average test scores. It was not expected to perform as well, since most factor data were for 1999. However, the factor weights and significance proved robust, and the model associated statistical variance as shown in Table 2-13.

Table 2-13
Year Comparisons for MCAS Grade 10 Math Test Scores

Year	R^2
1998	.83
1999	.80
2000	.77

Sources of data: Appendix 3, Appendix 4, Statistica model

The residuals in Table 2-11 suggest systematic contributions to test score averages at some schools that might have been produced by unusual efforts. Probable outliers for the predictive model with correspondingly large year-to-year average score changes had a strongly positive bias: those for Boston Latin, Latin Academy and Swampscott.

Model behavior for metropolitan Boston schools may be biased by special characteristics of City of Boston schools, because the students who score well on school-based standard tests are selected for admission to the three exam schools. (Note 9) Current data also attribute average Boston per-capita income equally to all school districts instead of adjusting by districts or census tracts. The Boston cross-enrollment and busing programs would complicate an income analysis. Behavior of the three-factor model of Table 2-10 was examined for 1999 tenth-grade MCAS mathematics test scores, considering only schools outside the City of Boston. These 34 schools had a 1999 total population per grade of about 10,200 students out of 13,730 for all schools considered in these studies.

When applied only to schools outside the City of Boston, the three-factor model of Table 2-10 showed significance at the $p<.05$ level for "Percent limited English proficiency" and "Per-capita community income (1989)" but not for "Percent Asian or Pacific Islander." A two-factor model based on the first two of these factors, with schools weighted by numbers of test participants, associates 86 percent of the statistical variance. Factor weights became as shown in Table 2-14.

Table 2-14
2-Factor Model for 1999 MCAS Grade 10 Math Test Scores

Factor	Coefficient	Standard Error
Intercept, for all factors zero	201.5	2.934

E. Percent limited English proficiency	-0.325	0.136
H. Per-capita community income (1989)	1.307	0.126

Sources of data: Appendix 3, Appendix 4, Statistica model

Several trial factors were added individually to the two-factor model of Table 2-14, but none showed statistical significance at the p<.05 level. The statistical variance associated when each trial factor was added to this two-factor model is shown in Table 2-15.

Table 2-15
Factor Comparison for 1999 MCAS Grade 10 Math Test Scores

R^2	Trial Factor Added
.86	None
.88	**A**. Population per grade
.88	**B**. Percent African American
.86	**C**. Percent Asian or Pacific Islander
.86	**D**. Percent Hispanic / Latino
.87	**F**. Percent free or reduced price lunch
.87	**G**. Percent reduction, grades 9+10 to 11+12
.87	Percent special education
.86	Per-capita property value, 1998, $000s
.88	Spending, regular education, $000s

Sources of data: Appendix 3, Appendix 4, Statistica model

The last three trial factors in Table 2-15 are districtaverages; the last and third from last are for all schools in all grades. Three districts, Lynn, Newton and Quincy, each operate two academic high schools which will not be distinguished by these factors. Outside the City of Boston, only "community income" and "limited English proficiency" are significant contributors to 1999 tenth-grade MCAS mathematics test scores; their indicators are effective predictors.

A one-factor model, using only "Per-capita community income (1989)," performed almost as well as any combination of factors shown in Table 2-15, associating 84 percent of the variance. The factor weight is in Table 2-16.

Table 2-16
1-Factor Model for 1999 MCAS Grade 10 Math Test Scores

Factor	Coefficient	Standard Error
Intercept, for all factors zero	197.0	2.395
H. Per-capita community income (1989)	1.465	0.114

Sources of data: Appendix 3, Appendix 4, Statistica model

"Percent limited English proficiency" is much less effective in a one-factor model, associating only 38 percent of the variance. Community income appears to be the dominant factor associated with these test scores. The 1999 adjusted average tenth-grade MCAS mathematics test scores by school, plus residuals from the two-factor and one-factor models for 1999, shown in Table 2-14 and Table 2-16, with standard error estimates for each, are in Table 2-17.

Table 2-17
Residuals for MCAS Grade 10 Math Test Scores, 1,2-Factor Models

City or Town	High School	Average Score	Std. Error	2-factor Residual	Std. Error	1-factor Residual	Std. Error
Arlington	Arlington	234	2.1	4.7	2.2	5.6	2.2
Belmont	Belmont	243	2.2	6.6	2.4	6.8	2.4
Braintree	Braintree	228	1.9	2.3	2.1	3.7	2.0
Brookline	Brookline	240	1.6	1.4	2.0	0.5	2.0
Cambridge	Rindge& Latin*	220	1.6	-5.0	1.8	-6.1	1.8
Chelsea	Chelsea	216	2.3	4.3	2.7	2.1	2.6
Dedham	Dedham	227	2.5	1.1	2.6	2.1	2.6
Everett	Everett*	221	1.9	2.6	2.1	3.2	2.2
Lexington	Lexington	238	1.7	-3.3	2.2	-4.0	2.2
Lynn	Classical	216	2.0	2.2	2.4	-0.1	2.2
Lynn	English	213	2.1	1.0	2.9	-3.1	2.4
Malden	Malden	221	2.0	2.2	2.2	0.9	2.1
Marblehead	Marblehead	232	2.7	-9.2	3.0	-9.8	3.1
Medford	Medford*	221	2.2	-1.7	2.3	-0.8	2.3
Melrose	Melrose	226	2.1	-1.7	2.3	-0.6	2.2
Milton	Milton	228	2.3	-2.6	2.4	-1.9	2.4
Newton	North*	239	1.5	0.6	1.9	-0.2	1.9
Newton	South	242	2.0	3.4	2.3	2.8	2.3
Peabody	Veterans*	220	1.8	-3.1	2.0	-1.9	1.9
Quincy	North Quincy	227	1.9	2.7	2.2	4.5	2.0
Quincy	Quincy*	212	2.3	-9.4	2.5	-10.5	2.5
Revere	Revere*	218	1.9	-1.3	2.1	-0.5	2.1
Salem	Salem*	220	2.1	-0.4	2.3	-0.6	2.3
Saugus	Saugus	226	2.2	1.2	2.4	3.0	2.3
Somerville	Somerville*	216	1.8	-0.4	2.3	-3.2	2.0
Stoneham	Stoneham	227	2.5	1.7	2.7	3.3	2.6
Swampscott	Swampscott	240	2.6	5.5	2.7	5.6	2.8
Wakefield	Memorial	227	2.2	0.7	2.4	2.2	2.3
Waltham	Waltham*	220	1.8	-2.3	2.0	-1.6	2.0

Watertown	Watertown	231	2.6	3.8	2.7	4.2	2.7
Weymouth	Weymouth*	222	1.5	-3.5	1.8	-1.9	1.7
Winchester	Winchester	243	2.3	1.5	2.7	1.2	2.7
Winthrop	Winthrop*	223	3.0	-1.3	3.2	-0.1	3.1
Woburn	Woburn	226	2.0	1.3	2.1	2.4	2.1

* school providing vocational education
Sources of data: Appendix 3, Appendix 4, Statistica model

In an attempt to improve accuracy of the model in Table 2-14, schools with residuals from the two-factor model for 1999 that were greater than two standard deviations were dropped, Belmont with positive residual and Cambridge Rindge & Latin, Marblehead High and Quincy High with negative residuals. The two-factor model for 1999 scores then produced the factor weights shown in Table 2-18.

Table 2-18
2-Factor Trial for 1999 MCAS Grade 10 Math Test Scores

Factor	Coefficient	Standard Error
Intercept, for all factors zero	200.6	2.188
E. Percent limited English proficiency	-0.216	0.101
H. Per-capita community income (1989)	1.357	0.095

Sources of data: Appendix 3, Appendix 4, Statistica model

Chi square for the two-factor model of Table 2-18 was 36.2 with 27 degrees of freedom (p=.11). The one-factor model of Table 2-16 was also evaluated for 1999 with the set of cases used in Table 2-18, producing the factor weights shown in Table 2-19.

Table 2-19
1-Factor Trial for 1999 MCAS Grade 10 Math Test Scores

Factor	Coefficient	Standard Error
Intercept, for all factors zero	197.6	1.793
H. Per-capita community income (1989)	1.463	0.087

Sources of data: Appendix 3, Appendix 4, Statistica model

Chi square for the one-factor model of Table 2-19 was 43.4 with 28 degrees of freedom (p=.03). However, the cases and models from Table 2-18 and Table 2-19 did not provide significant chi square probabilities for 1998 or 2000 scores; all other attempts to improve estimation by removing outliers also proved unstable.

Residuals from the two-factor model of Table 2-14 for schools outside the City of Boston are strongly autocorrelated. The coefficient was .45 between 1998 and 1999 and .67 between 1999 and 2000. Scatterplots for successive years are shown in Figure 2-4 and Figure 2-5, comparing 1998 and 2000 residuals with 1999 residuals:

Figure 2-4: MCAS Grade 10 Math Test Score Residuals, 1998 versus 1999

.

.

Figure 2-5: MCAS Grade 10 Math Test Score Residuals, 2000 versus 1999

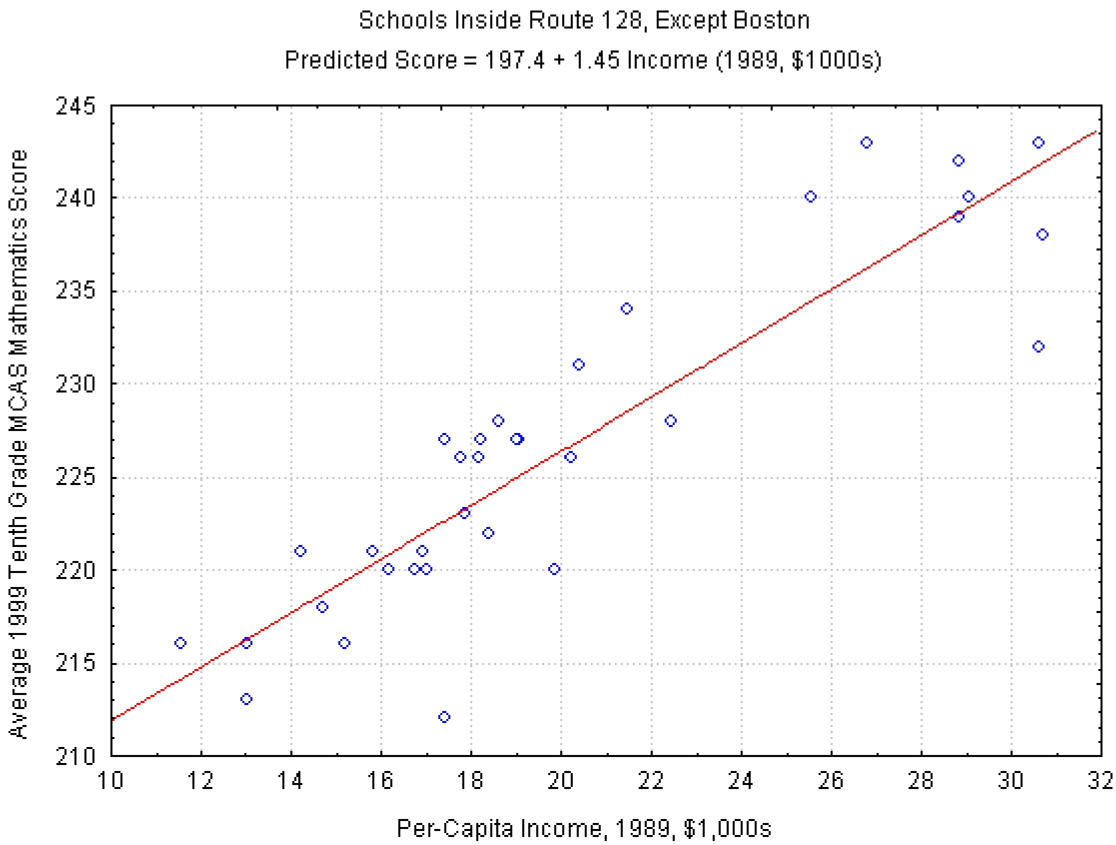
To the extent that a trend can be observed by applying the model of Table 2-14 to different years, it appears that schools scoring higher than predicted tend to increase scores in successive years, and schools scoring lower than predicted tend to decrease scores. Departures from model predictions are not all random. The two spans of years available for analysis suggest a systematic trend that could stratify high-scoring and low-scoring schools.

C. Observations

The trends and effects studies presented in Section 2A and Section 2B show how studies of these types tend to yield results with low statistical significance unless "something unusual" is going on. In the cases of the 2000 tenth-grade mathematics test and of the high scores and large increases at Boston Latin, Latin Academy and Swampscott, the studies do indicate "something unusual," although they cannot tell what it is. They also illustrate that commonly published "league tables" of scores strongly reflect social factors associated with school populations, not factors clearly associated with school performance. The effects study shows a robust, positive correlation of scores with household incomes, plus a smaller, negative correlation with limited English proficiency. Factor weights and statistical significance for other factors considered are small. School-based data are not currently published for several potentially significant factors, such as class sizes in elementary grades, mathematics course enrollments and levels of teacher preparation.

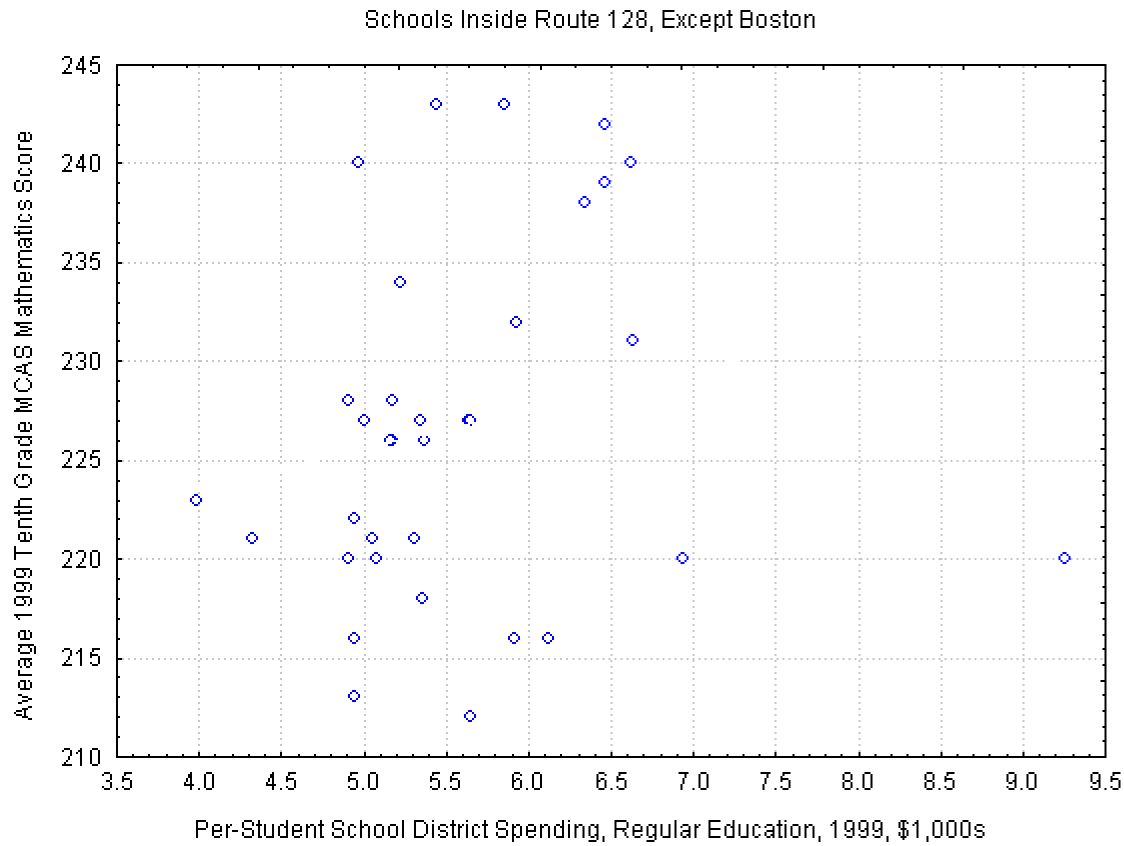
By far the strongest factor in predicting tenth-grade MCAS mathematics test scores is "Per-capita community income (1989)." For the schools outside the City of Boston this factor alone performed nearly as well as all available factors combined, associating 84 percent of the variance compared with 88 percent when all available factors were used. The scatterplot in Figure 2-6 shows 1999 average scores for these schools versus 1989 per-capita community income:

Figure 2-6: 1999 MCAS Grade 10 Math Test Scores versus Community Income



In Figure 2-6, the relation between school-averaged test scores and per-capita community income looks linear over an income range of about 2½ to 1 for this set of schools. There is no obvious threshold or saturation behavior. As it happens, spending on regular education programs in school districts also varies over a range of about 2½ to 1 for those schools (Mass. DoE, 2000f). However, there is only a weak relationship between 1999 spending on regular education programs in those school districts and their 1999 tenth-grade MCAS mathematics test scores, associating only about 3 percent of the variance and not statistically significant at a $p<.05$ level, as shown in the scatterplot of Figure 2-7.

Figure 2-7: 1999 MCAS Grade 10 Math Test Scores versus School Spending



Although a tendency for scores on school-based standard tests to rise with incomes has been recognized in the US for more than 70 years (Bolon, 2000), there has been relatively little research on this phenomenon in ordinary income ranges, compared with the attention given to associations between test scores and conditions of poverty or race. The strong, apparently linear association of average test scores with community incomes shown here, as contrasted with their weak association with school spending, calls for investigation but is beyond the scope of this report. It seems likely that community incomes are providing something beyond what school programs provide, but if so we cannot tell from these data what it might be. Perhaps this effect should not be surprising, since most students spend more than three-fourths their waking hours outside school.

The factor "Percent limited English proficiency" was the second strongest influence on predicted test scores. (Note 10) Previous effects studies of MCAS scores might not have shown this if they failed to utilize the entire variety of school-associated data available, including "limited English proficiency," "racial" or "ethnic," and "free or reduced price lunch" student categories. A hypothesis inviting study is that students classified with "limited English proficiency" might receive mathematics instruction less relevant to curriculum tested by MCAS than other students. Another hypothesis is that strengthening English language instruction for "limited English proficient" students might improve their MCAS mathematics test scores, provided all current efforts to teach mathematics and other subjects are maintained. While statistical associations suggest these conjectures, only observations and experiments could prove or disprove them.

Factors of "Percent African American" and "Percent Hispanic / Latino" did not make significant contributions to school-averaged tenth-grade MCAS mathematics test scores after other factors were introduced. Although "Percent Asian or Pacific Islander" retains statistical significance in certain models for some years, it is a much weaker factor than "Per-capita community income (1989)" or "limited English proficiency." The latter factors, and not "racial" or "ethnic" percentages, provide by far the strongest statistical associations with school-averaged tenth-grade MCAS mathematics test scores in metropolitan Boston. (Note 11)

Statistical significance of test-based ratings has been the principal focus of the studies. Some studies developing or using such ratings do not provide an analysis of variability, without which significance cannot be determined; or they estimate variability from single test session reliability measurements, which were shown to be optimistic (for example, Gaudet, 2001, and Tuerck, 2001a). The trends study in Section 2A of this report provides evidence derived from tenth-grade MCAS mathematics test scores for larger variability estimates.

D. Summary Analysis

My summary analysis is based on a one-factor model for metropolitan Boston communities that each operate only a single academic high school, weighted by numbers of students tested. The effects study showed that "Per-capita community income (1989)" was the dominant factor in predicting 1999 school-averaged tenth-grade MCAS mathematics test scores. All other factors made only small contributions to predictions with much lower significance. From published data, community income could not be estimated reliably for each of the multiple academic high schools in Boston, Lynn, Newton and Quincy. Reduction in scope leaves 28 high schools in the same number of communities, with a total of about 8,200 students per grade recorded for 1999. Estimates of uncertainties in school-averaged test scores are based on findings of the trends study, which showed year-to-year variability of school-averaged scores several times greater than the variability implied by conventional test reliability measurements. Plots of results include uncertainty intervals (sometimes called "error bars") equivalent to +/- 1.4 estimated standard errors. When item intervals do not overlap, when standard errors have been accurately estimated, and when items are uncorrelated, then differences between items are significant at the $p<.05$ level. The one-factor model for 1999 with this set of cases associated 82 percent of the variance and produced the factor weights shown in Table 2-20.

Table 2-20
Community Income Model for 1999 MCAS Grade 10 Math Test Scores

Factor	Coefficient	Standard Error
Intercept, for all factors zero	198.4	2.667
Per-capita community income (1989)	1.397	0.129

Sources of data: Appendix 3, Appendix 4, Statistica model

The 1999 adjusted average tenth-grade MCAS mathematics test scores by school, plus residuals from the foregoing one-factor model for 1999, with standard error estimates for each, are shown in Table 2-21.

Table 2-21
Residuals for MCAS Grade 10 Math Test Scores, Income Model

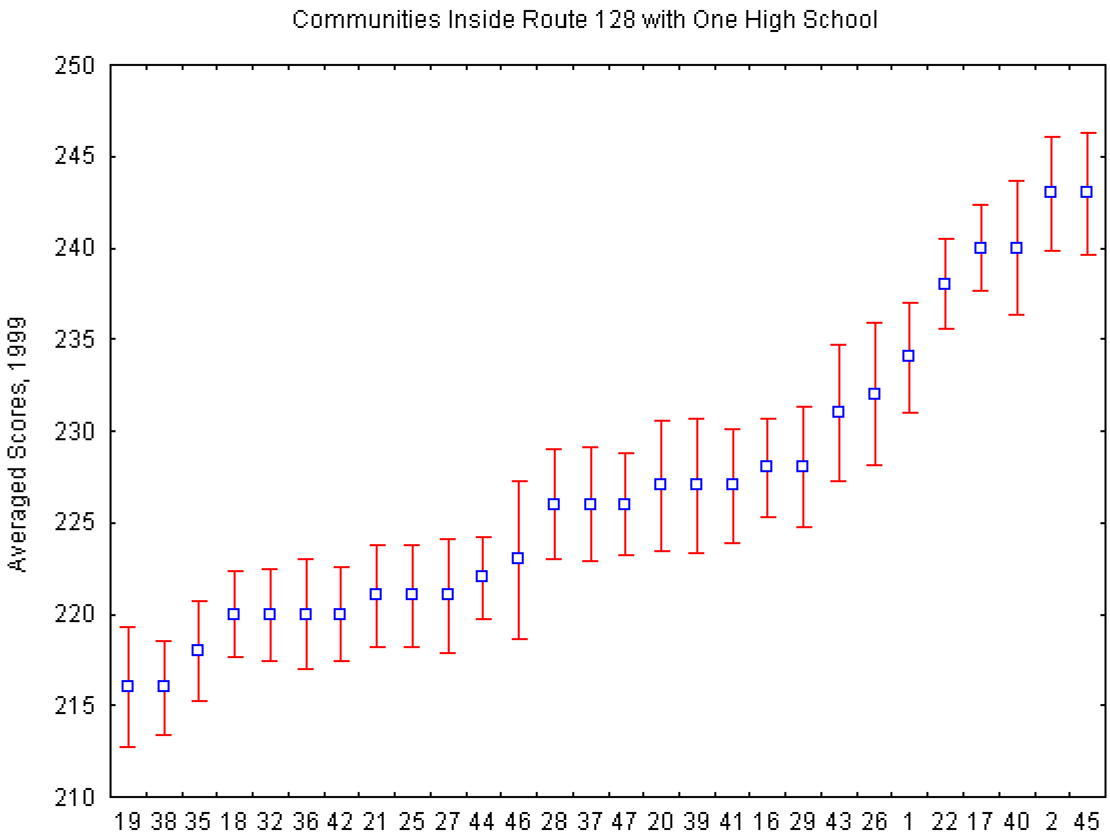
City or Town	Average Score	Std. Error	1-factor Residual	Std. Error
Arlington	234	2.1	5.6	2.2
Belmont	243	2.2	7.1	2.5
Braintree	228	1.9	3.5	2.0
Brookline	240	1.6	1.0	2.1
Cambridge*	220	1.6	-6.2	1.8
Chelsea	216	2.3	1.4	2.6
Dedham	227	2.5	2.0	2.6
Everett*	221	1.9	2.7	2.2
Lexington	238	1.7	-3.4	2.3
Malden	221	2.0	0.5	2.2
Marblehead	232	2.7	-9.2	3.2
Medford*	221	2.2	-1.1	2.3
Melrose	226	2.1	-0.7	2.2
Milton	228	2.3	-1.8	2.4
Peabody*	220	1.8	-2.2	1.9
Revere*	218	1.9	-1.0	2.1
Salem*	220	2.1	-1.0	2.3
Saugus	226	2.2	2.7	2.3
Somerville*	216	1.8	-3.6	2.0
Stoneham	227	2.5	3.1	2.7
Swampscott	240	2.6	5.8	2.8
Wakefield	227	2.2	2.0	2.3
Waltham*	220	1.8	-1.9	2.0
Watertown	231	2.6	4.1	2.7
Weymouth*	222	1.5	-2.1	1.7
Winchester	243	2.3	1.8	2.8
Winthrop*	223	3.0	-0.4	3.1
Woburn	226	2.0	2.2	2.1

* school providing vocational education

Sources of data: Appendix 3, Appendix 4, Statistica model

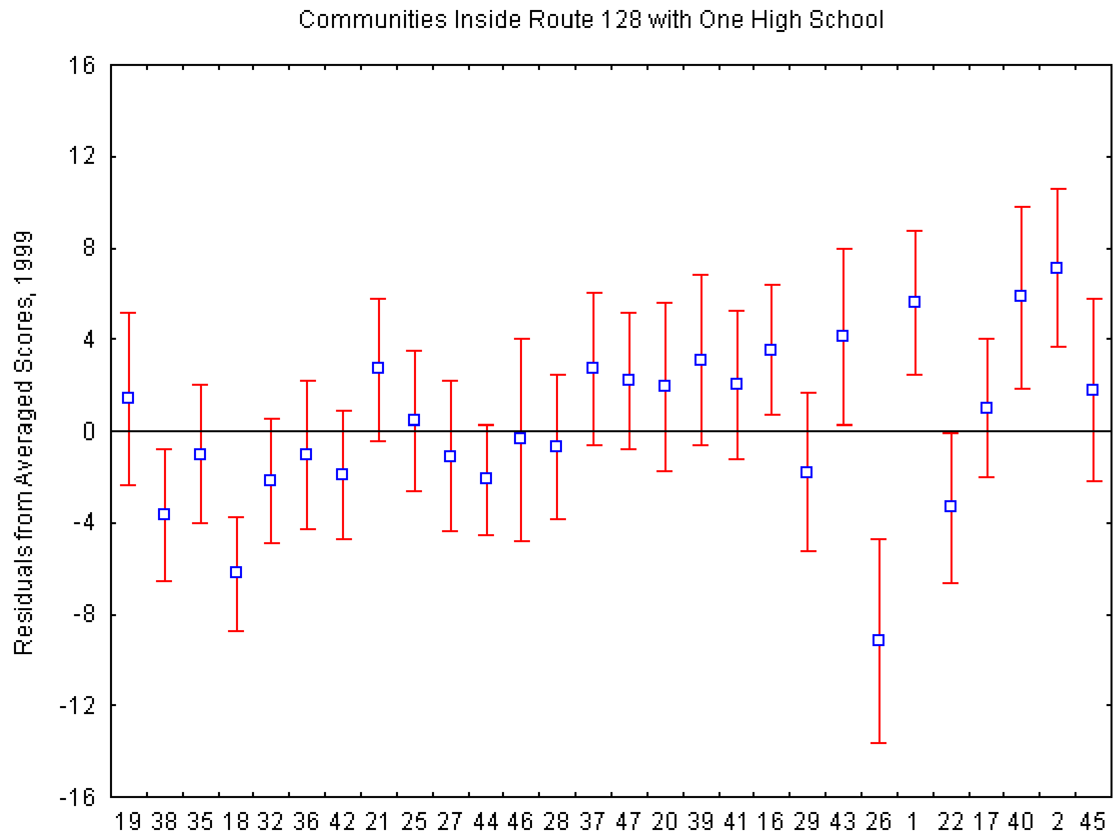
The plot in Figure 2-8 shows school-averaged adjusted scores on the 1999 tenth-grade MCAS mathematics test and the corresponding uncertainty intervals (or "error bars"). In this and the next two plots, the 28 schools considered in this summary analysis have been rank-ordered from the lowest to the highest average scores on the 1999 tenth-grade MCAS mathematics test.

Figure 2-8: Average 1999 MCAS Grade 10 Math Test Scores by School



The plot in Figure 2-9 shows residuals for school-averaged 1999 tenth-grade MCAS mathematics test scores, the differences left after subtracting away predictions calculated from the factor "Per-capita community income (1989)."

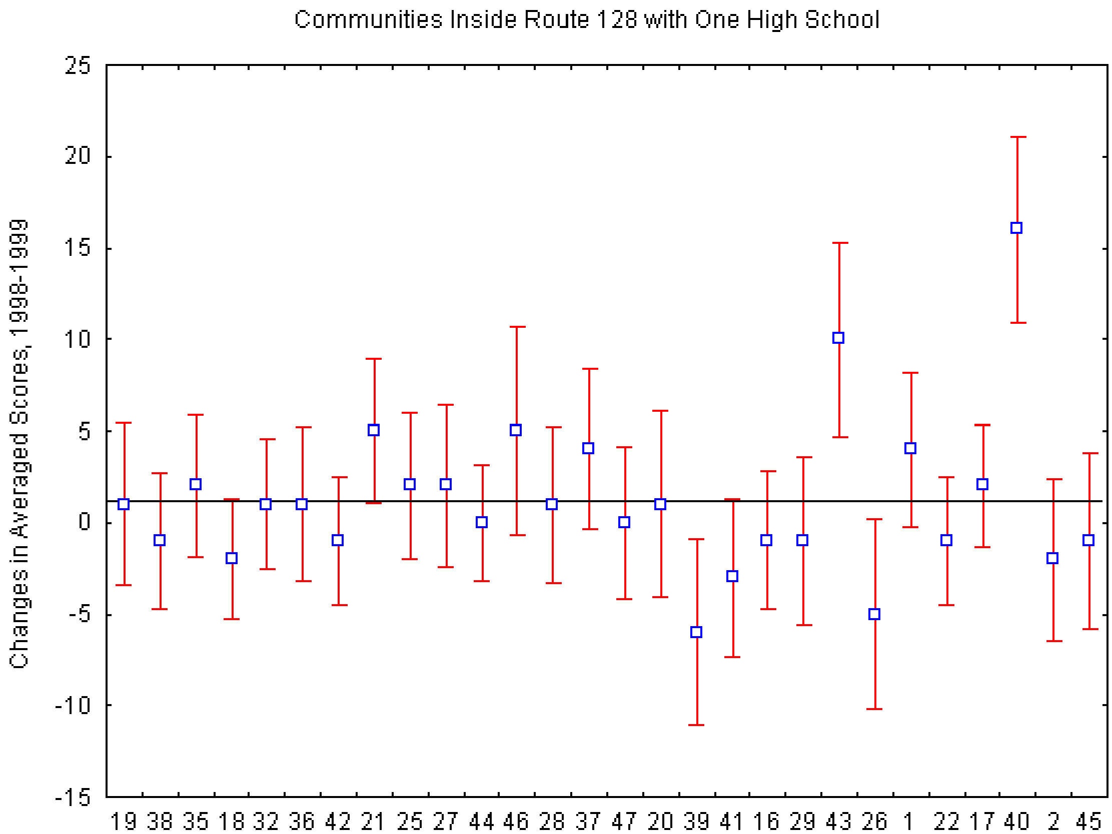
Figure 2-9: Residuals of 1999 MCAS Grade 10 Math Test Scores by School



The picture in the plot of average scores of Figure 2-8, with significant separations between high-scoring and low-scoring schools, is shown by the residuals plot of Figure 2-9 to be associated largely with differences in community income. Chi square for the residuals distribution is 63.2 with 27 degrees of freedom. After subtracting predictions based on community income, residuals of average scores settle to just a little more than statistical noise. Only five or six schools can be reliably distinguished other than by community income.

The last plot, Figure 2-10, shows the changes in school-averaged 1999 tenth-grade MCAS mathematics test scores from the same schools' average scores in 1998.

Figure 2-10: Changes in MCAS Grade 10 Math Test Scores by School, 1998-1999



There was low statistical significance in year-to-year changes of the school-averaged adjusted scores on tenth-grade MCAS mathematics tests between 1998 and 1999. The distribution of changes about an average change of +1.2 scale points, as shown in Figure 2-10, fits a normal distribution with weighted chi square of 22.7 for 26 degrees of freedom (p = .65). Despite the possibility of little significance in score changes, such changes are criteria by which the Massachusetts Department of Education has begun to rate school Performance (Note 12) as "Failed to Meet Expectations," "Approached Expectations," "Met Expectations" or "Exceeded Expectations." There are severe penalties, including state closure or seizure, for schools with low scores and ratings of "Failed to Meet Expectations."

The statistical significance of differences in school-averaged scores, aside from their reflection of community income, may not be not enough to compare schools reliably and may not be enough to evaluate short-term changes in teaching and learning. Whatever aspects of school performance these scores might measure can be lost in the fluctuations for a particular school over any one year or few years. Only averages and trends in scores over several years would be likely to yield statistically useful information.

Some observers question whether it is realistic to expect standardized tests to yield significant information comparing school performance, even over a period of years (for example, Rowe, 1999a, and Rowe, et al., 1999b). They argue that variations in student achievement are commonly greater within schools than between schools. Others contend that adaptive and defensive behavior encouraged by "high-stakes" political environments grossly distorts outcomes from all types of educational assessment and robs them of meaning (for example, Sacks, 1999, and Hayman, 1998). It is important to point out that significance attributed to results in these studies is purely statistical. Neither these studies nor any others known to the author have shown that MCAS test scores have practical significance, in the sense of predicting success in adult activities to any greater degree than could be done with knowledge of student backgrounds.

Section 3: Results

A. Opportunities and Questions

In future work, it may be helpful to examine categories of metropolitan Boston schools that were excluded from or specially identified in these studies:

- Vocational, technical and agricultural schools
- Academic schools that also provide vocational education
- Pilot schools, charter schools, specialty schools and small schools

Other questions might be answered by extending data coverage:

- Do other test scores yield similar results?
- Do small class sizes have significant effects?
- Does teacher preparation have significant effects?
- Do math course enrollments have significant effects? (Note 13)
- Can other factors be found that increase significance?
- Are the same effects found in other Massachusetts schools?
- Are similar effects observed in elementary and middle school grades?
- Are the same or different patterns observed in other US metropolitan areas?
- Can income factors be estimated for communities with multiple high schools?
- If individual data were available, would multi-level analysis show different results?

The strong, linear relation found between school-averaged tenth-grade MCAS mathematics test scores in a relatively calm year and community income in a previous year, within ordinary ranges of incomes, leads to several questions. Does community income primarily determine educational achievement, regardless of school performance? Do MCAS and similar tests measure skills and knowledge that are acquired in schools, or do they measure skills and knowledge that are largely acquired outside schools? Do communities with substantially different incomes have substantially different expectations for student performance on MCAS and similar tests? Is current community income also strongly correlated with test scores? After a turbulent period of local efforts to raise scores, will the correlation between income and scores remain as strong? Exploring and understanding some of these issues will require a different approach.

B. Conclusions

Community income has been found strongly correlated with tenth-grade MCAS mathematics test scores and associated more than 80 percent of the variance in school-averaged 1999 scores for a sample of Boston-area communities. The influence of community income was robust against several sets of model variables and cases.

Community income swamped the influence of the other social and school factors examined. Once community income was included in models, other factors--including percentages of students in disadvantaged populations, percentages receiving special education, percentages eligible for free or reduced price lunch, percentages with limited

English proficiency, school sizes, school spending levels, and property values--all failed to associate substantial additional variance.

Large uncertainties in residuals of school-averaged scores, after subtracting predictions based on community income, tend to make the scores ineffective for rating performance of schools. Large uncertainties in year-to-year score changes tend to make the score changes ineffective for measuring performance trends. In their present state, considered as a means to rate the performance of public schools, tenth-grade MCAS mathematics tests mainly appear to provide a complex and expensive way to estimate community income.

Appendix 1: Education Reform in Massachusetts

Massachusetts has experienced many education experiments and reforms over the past few centuries. (Note 14) During the 1990s Massachusetts education reform was driven by the McDuffy school finance lawsuit (McDuffy, 1993), originally filed in 1978. In its McDuffy decision, the Massachusetts Supreme Judicial Court said that Massachusetts funding disparities harmed the quality of education for some students, denying them education to which they were constitutionally entitled. This June, 1993, decision was widely anticipated. The Massachusetts Education Reform Act of 1993 (Education Reform, 1993) was signed less than a week after the decision was released.

A group called the Massachusetts Business Alliance for Education, (Note 15) organized in 1988 and led by the late John C. (Jack) Rennie, then CEO of the former Pacer Infotec, Inc., of Burlington, MA, (now the AverStar division of Titan Corp., San Diego, CA), and S. Paul Reville, then director of the Worcester Public Education Fund, wrote the reform bill sponsored by the Education Committee of the legislature. In 1991 the Business Alliance produced a document entitled *Every Child a Winner*. (Note 16) A story from the May 2, 1993, Northwest edition of the *Boston Globe* quoted former Rep. Mark Roosevelt, then House Education Committee Chair, as saying that the House education reform bill then pending "is essentially [the Business Alliance document]." In a publication of MassINC, Rennie is quoted as saying, "We bought change" (Walser, 1997). Most of this work was carried out in secret. As late as December, 1992, then Lt. Gov. Cellucci was calling on the Education Committee chairs, Sen. Thomas Birmingham of Chelsea and Rep. Roosevelt of Beacon Hill, to disclose their bill (Howe, 1992). Almost all the controversy generated by this legislation focused on its funding formulas. Until 1993, the public had hardly any knowledge of its sweeping changes in school policy and regulation. The following newspaper report was printed December 23, 1992 (Overdue, 1992):

> "The bill also calls for higher student achievement and curriculum standards."

This was the most thorough description in mainstream news media from 1988 through 1992. The bill was released in an emergency legislative session of January 4-5, 1993, but it failed to pass.

Soon after the bill became public, education and public interest groups began to react. As reported in the *Boston Globe* on January 26, 1993, a coalition headed by Stephen Bing of the Massachusetts Advocacy Center predicted major problems with the legislation, including these (Ribadeneira, 1993):

- The reform bill will institutionalize unfair teaching practices such as using tests to track students into different ability levels.

- By substituting different certificates in place of the high school diploma, the bill will contribute to the dropout problem rather than ameliorate it.

- The legislation provides no mechanism for meaningful participation by parents or students in the development of a remedial education plan nor any opportunity to contest an inadequate plan.

Such objections were ignored. Neither the mainstream news media nor the Great and General Court gave these or other educationally oriented issues further attention in 1993.

Rep. Thomas Finneran of Mattapan, then chair of the House Ways and Means Committee, secretly inserted anti-abortion provisions in the bill, provoking a storm of protest (Howe, 1993a). Other House controversy centered on a salary cap for teachers, which was removed. Proposals for "school choice," charter schools and gambling revenues became the focus of activity in the Senate. The bill quickly became a hodge-podge of added provisions with no coordination. Many observers became skeptical about overall benefit. Geoffrey Beckwith of the Massachusetts Municipal Association was quoted as saying, "It certainly doesn't appear at this time that this bill will bring about any fundamental reform" (Howe, 1993b). In February, 1993, the Supreme Judicial Court heard testimony in the McDuffy case. In March, former Rep. Roosevelt began a (losing) campaign for Governor (Lehigh, 1993, Howe, 1993c). In April, the Edison Project, a business corporation, announced interest in privatizing Massachusetts schools (Nealon, 1993). By May, an impasse over "school choice" had developed, the then Senate President William Bulger of Boston demanding it and the then House Speaker Charles Flaherty of Cambridge rejecting it. At the time, the Business Alliance opposed the "school choice" and charter school amendments (Taylor, 1993). However, another business group calling itself "CEOs for Fundamental Change in Education" had appeared, dominated by banking and large business interests and supported by the Pioneer Institute. It was actively promoting charter schools and "school choice" through the Massachusetts Senate (Vennochi, 1993). After compromising with limits and delays on "school choice" and charter schools, the House passed the bill through second reading June 2 and the Senate passed it June 3. The Supreme Judicial Court released its McDuffy case decision June 15. Former Gov. William Weld signed the Education Reform Act on June 18, 1993.

In seven years under the Education Reform Act, state aid to Massachusetts public schools has grown from $1.3 billion to $3.0 billion per year, almost all the increase going to communities with low household incomes. For example, Holyoke, a low-income community, now receives over 90 percent of its school funding from the state, while Brookline, a high-income community, receives only about 10 percent (Mass. DoE, 2000e). In 1992, Holyoke spent less than 75 percent as much per student as Brookline, but now it spends about 95 percent of what Brookline does (School reform, 1992). Still, the Act has tended to provide more of a windfall for Holyoke's taxpayers than for its public school students.

Besides setting state commitments to equalize school funding, the Education Reform Act made many changes to Massachusetts education policy and regulations, including the following, as described by the Business Alliance (Taylor, 1993):

- New goals, standards and indicators of performance for schools, students and teachers

- Financial rewards to teachers and schools that excel
- Decentralized authority, limiting school committees to policy-making and oversight, making CEOs of superintendents, and giving hiring and firing power to principals
- Preschool for all 3- and 4-year-olds
- Expanded professional development for teachers

More than seven years later, some of these changes are only starting to be implemented.

Before 1996, the Massachusetts Board of Education regarded test scores as only one component of school accountability. In a 1993 policy advisory cited by Wheelock, 1999, the Board warned that an accountability system based primarily on test scores would be likely to produce harmful long-term consequences, including: exclusion of weaker students from the assessed pool of students; lowered morale among teachers and students; the loss of experienced educators from schools enrolling many disadvantaged students; distortion of instruction and curriculum to reflect test content and format; cheating and corruption of test scores.

Nevertheless the Board began development of testing programs to satisfy the provisions of the Education Reform Act, which eventually became MCAS. It appointed committees of educators and parents to help insure that tests were meaningful, fair and free from overt forms of bias (French, 1998). From 1993 through 1996, Massachusetts invested more than $2 million to support education reform study groups seeking ways to set high expectations for students (Antonucci, 1997a).

By March of 1996 the Department of Education had completed a Common Core of Learning (Mass. DoE, 1994), released six of seven planned curriculum frameworks based on it, and begun the development of MCAS based on the frameworks. In addition, it had announced plans (Mass. DoE, 1996a) to: award grants to school districts for assessment activities such as portfolio development; hold statewide conferences on local assessment strategies; publish examples of student work that meet the statewide standards so that districts have a model of what to strive for; develop a bank of assessment exercises linked to the curriculum frameworks for use by classroom teachers. All of these satisfy or support provisions of the Education Reform Act.

Development of MCAS took a sharp turn away from public participation (Mass. DoE, 1996b) after the appointment of John Silber as chair of the Massachusetts Board of Education in November, 1995 (Pawlack-Seaman, 1996). In August, 1996, Silber, former president of Boston University and an unsuccessful candidate for Governor, working with then Gov. Weld, engineered replacement of the 17-member Board of Education, including four African-Americans and Latinos, with a 9-member board, including several with ties to school privatization and charter schools, only one African-American and no Latinos (Jackson, 1996, Wong, 1996). At his first meeting with the Board, Silber said the Education Reform Act's underlying principle that all students are capable of learning at high levels was "rubbish" (Future, 1996). Responding to a demand for his resignation in 1997 he commented, "Some of the things that pass for learning disabilities used to be called stupidity" (Pawlack-Seaman, 1997). Soon after the Board replacement, the committees of educators and parents that had been formed to oversee curriculum frameworks, test development and other education reforms were disbanded (Antonucci, 1997b).

In December, 1996, Silber proposed a two-track system (Avenoso, 1996) with a general diploma awarded for passing the GED, a test introduced during World War II by Everett F. Lindquist, developer of the Iowa test series, and now administered by the American Council on Education. An honors diploma would be awarded for high scores on the Massachusetts test series. Silber was forced to abandon the plan in January, 1997, when his personally chosen Board of Education refused to support it (Leung, 1997). However, a legacy of Silber's proposal remains, the view that MCAS should be aimed at the exceptional student. In August, 1999, the Business Alliance revived the two-track concept (Still, 1999) with a proposal to award general diplomas to students who satisfy "essential requirements in English and math." The Business Alliance did not specify how this would be administered, and the Department of Education and Board of Education still oppose the concept. What they have done instead is to make a "competency determination" required by the Education Reform Act for a high-school diploma depend on achieving relatively low MCAS test scores, (Note 17) answering about 40 percent of the questions. (Note 18) A "certificate of mastery," as specified by the Act, is to be awarded for much higher scores, (Note 19) answering about 80 percent of the questions to achieve an "advanced" rating on one or more tests.

After the loss of two Education Commissioners in rapid succession (Battenfield and Pressley, 1999), Silber resigned during a struggle over a new Commissioner in March, 1999. The outcome of the controversy (Estrin, 1999) was replacement of Silber by James Peyser (see Peyser, 1996, and Peyser, 1998), head of the reactionary Pioneer Institute, tied to school voucher and privatization movements, and retention of the compliant acting Commissioner David Driscoll. Since the Silber era, MCAS development has been closely monitored by Board of Education member Abigail Thernstrom, a fellow of the Manhattan Institute, (Note 20) and hired consultant Sandra Stotsky (see Stotsky, 1999), a writer for the Fordham Foundation and now an Associate Commissioner of Education. Both of these right-wing foundations have supported forms of school privatization. Four rounds of MCAS tests have now been administered, in the spring of 1998, 1999, 2000 and 2001. The Board of Education has made the questions used in scoring available to the public, although they have not disclosed their standards for evaluating essay questions or all the details of their approach to computing scores.(Note 21) Students in religious-run and other private schools and students being taught at home are not required to take or pass MCAS tests. Bills have been filed but have not been enacted to include private schools in testing and to exclude charter schools. A system of "school accountability" has been defined by the Department of Education (Mass. DoE, 1999b). It is based entirely on MCAS scores, a violation of Education Reform Act requirements. (Note 22)

MCAS has been heavily promoted by a business-oriented group organized as Mass Insight Education and Research Institute, Inc., in Boston, founded in 1997 by registered Massachusetts lobbyist William H. Guenther, who is its president. (Note 23) Guenther is also involved with three other public relations organizations, Mass Insight Corp., in Cambridge, Opinion Dynamics Corp., in Cambridge, and New England Economic Project, in Walpole. Mass Insight Education and Research Institute is a non-profit corporation that coordinates several policy groups and has close relationships with business and education executives. Leaders of its "Campaign for Higher Standards" have included Gloria Larson, former Mass. Secretary of Economic Affairs, the late John C.

Rennie, former Chairman of the Massachusetts Business Alliance for Education and Vice-Chairman of AverStar, Inc., and Cathy Minehan, President of the Federal Reserve Bank of Boston. Leaders of its "Coalition for Higher Standards" include James Caradonio, Superintendent of Worcester Public Schools, and Thomas Payzant, Superintendent of Boston Public Schools. Its board of directors has included Maura Banta, Manager for External Programs at IBM Corporation, John Rennie, Abigail Thernstrom, Senior Fellow at the Manhattan Institute and member of the Massachusetts Board of Education, and Bruce Tobey, Mayor of Gloucester. Financial supporters of Mass Insight Education and Research Institute include BankBoston (now FleetBoston Financial), State Street Corp., Bell Atlantic (now Verizon), Boston Edison (now an NSTAR division), Liberty Mutual Group, PricewaterhouseCoopers, Goodwin, Procter & Hoar, AverStar, Inc. (now a division of Titan Corp.), Gorton's Seafoods, Hewlett-Packard, IBM and Intel.

Mass Insight publications promoting MCAS have been distributed to public schools through the Massachusetts Board of Education, (Note 24) and Mass Insight has received public funds for its services. Mass Insight has been cited in minutes of the Board of Education as a source of policy initiatives, (Note 25) including a proposal to use a score of 220 on tenth-grade language arts and mathematics tests as the initial "competency determination" for high-school graduation, which was adopted by the Board in November, 1999. Mass Insight presents a simple but misleading picture of MCAS, saying that it measures "skills that students will need after graduation—at college or on the job" (Why, 1999). No such significance has ever been demonstrated for MCAS or other state accountability tests.

Massachusetts schools are often castigated by newspapers and politicians as mediocre, (Note 26) but actually they are superior. In the October, 1999, *Boston Magazine,* Jon Marcus wrote:

> "According to assessment tests and other measures, Massachusetts schools are among the nation's best. Students here rank fourth nationally in reading, sixth in math, and eighth in science on the National Assessment of Educational Progress, administered by the US Department of Education. They scored higher this spring in reading than 69 percent of their peers across the country on the Iowa Test of Basic Skills; a third of the state's third graders were at the advanced level, compared to 19 percent nationwide. A Boston College correlation of NAEP results with international tests found that in eighth-grade science Massachusetts students performed as well as, or better than, their counterparts in 40 out of 41 other countries, including Germany and Japan; only kids in Singapore were rated higher. More students study algebra and upper-level math and science than the national average, and Massachusetts also has the fifth-lowest high school dropout rate, the nation's highest percentage of graduates who enroll in college, and the third-highest proportion of students who take the SAT. Massachusetts students' SAT results have risen steadily since 1994; last year, they were the highest in a decade." (From Marcus, 1999.)

Such a contrast between political bombast and educational reality has become common. Part of the long record of declining SAT scores in the 1960s and 1970s, for example, had a straightforward cause, the rapidly expanding number of students taking the tests, including many low-income students and students with lower grades who would not have taken them in prior years (Koretz, 1992, Berliner and Biddle, 1995, Berliner and Biddle, 1996). When education researchers looked at comparable groups of students, SAT scores were gradually rising during much of this period; unadjusted averages began to rise as the growth in the number of test takers slowed. With the gratuitous abuse regularly heaped on public schools during this time, few members of the public would have guessed that some of the real trends in scores were positive. Even now, many politicians and most news media find the actual results inconvenient; they prefer simple, strident bashing of public schools, uncomplicated by facts.

Many observers and columnists have commented on the complex language, mental tricks and obscure bits of knowledge found in MCAS questions (see, for example, Vaishnav, 2000, and Kohn, 2000). How elitist are the MCAS tests? One way to look at this is to ask the fraction of questions that must be answered to pass them and the fraction of students who cannot do this. Table A1-1 compares the current graduation level tests in Massachusetts (Note 27) (10th grade MCAS) with those in New York (Note 28) (the revised Regents series) and Texas (Note 29) (the TAAS series).

Table A1-1
Comparison of State Achievement Tests

State	Typical percent of questions to pass	Typical percent of students failing (Note 30)
Texas	**70**	**20**
New York	**55**	**20**
Massachusetts	**40**	**50**

Source of data: see text, Appendix 1

Massachusetts, with by far the lowest passing score, has by far the highest rate of failure. Yet year after year, nationwide measures of academic performance rate Massachusetts students well ahead of those in New York and Texas. (Note 31) Passing an MCAS test says little about the education imparted through many years of schooling. On an MCAS tenth-grade mathematics test, the difference between passing and failing can be getting 24 questions right rather than 23 (see Mass. DoE, 1999d, and Mass. DoE, 1999c).

A recent study performed by Catherine Horn and others at Boston College (Horn, et al., 2000) showed that a barely passing score on the tenth-grade MCAS math test was approximately equivalent to the 50th percentile score for the PSAT math test. Students taking the PSAT are aiming for college. Many are taking the test as part of applying for National Merit and other scholarships; they tend to be good students. Therefore it should not be surprising when half or more of the general student population may "fail" the current tenth-grade MCAS math test.

A large share of MCAS test questions is aimed at students with exceptional skills and knowledge rather than at typical students. If Massachusetts designed tests to measure competence rather than mastery, it would be setting much higher passing percentages. If Massachusetts genuinely cared about assessing student skills and knowledge, it would satisfy Education Reform Act requirements (Note 32) calling for a "variety of assessment instruments," including "consideration of work samples, projects and portfolios," facilitating "authentic and direct gauges of student performance," and it would provide for circumstances of special education students, students entering the public schools from households that speak a first language other than standard English, (Note 33) and students whose immediate aims focus on employment rather than higher education.

The Massachusetts Board of Education has ample access to information of this sort and has received many recommendations to improve its practices and make its system of assessments more realistic and fair. It has had more than $25 million to spend on developing MCAS (Szechenyi, 1998). It is also well aware that "high-stakes" testing systems in other states have sharply narrowed the school curriculum (Note 34) and increased the population of school dropouts, (Note 35) who are likely to be eligible only for the "McJobs" of the future. So far, however, members of the Massachusetts Board of Education remain rigid, programmatic and hostile to facts that do not support their policies. Their attitude does not originate from lack of information or resources.

MCAS, like the other "achievement tests" used in state accountability systems, has never been shown to predict success in adult life to any greater degree than could be done with a knowledge of student backgrounds. Instead of trying to show practical significance the Board assumes it, in proposing to use this test as the sole state criterion to deny high-school diplomas and state college eligibility to low-scoring students, making it difficult for them to find responsible jobs and other forms of advancement. Students from households that already have the least suffer the most from such a system, tending to widen an economic gap between haves and have-nots in our society, already among the greatest of the industrial nations.

Appendix 2: Massachusetts Vocational Schools

Massachusetts municipal school districts support either jointly or individually more than 30 "vocational," "vocational-technical," "technical," "agricultural-technical" and "agricultural" high schools, (Note 36) most of which the Department of Education recognizes as separate school districts. Table A2-1 includes the 29 vocational schools that are now operated as separate school districts plus the "technical" high school operated by the City of Boston. (Note 37)

Table A2-1
1999 MCAS Grade 10 Math Test Scores for Vocational Schools

Vocational School	1999 MCAS math 10	
	Average	Number
Assabet Valley Voc. High	211	186
Bay Path Reg. Voc. Tech.	210	245
Blackstone Valley School	213	209
Blue Hills Reg. Voc. Tech.	213	215
Bristol County Agr. High	212	100
Bristol Plymouth Voc. Tech.	208	207
Cape Cod Reg. Voc. Tech. High	214	171
Charles McCann Voc. Tech.	217	105

Diman Reg. Voc. Tech. High	208	312
Essex Agr. & Tech. Inst.	212	80
Franklin County Tech.	214	115
Gr. Lowell Reg. Voc. Tech.	205	466
Gr. New Bedford Voc. Tech.	208	497
Greater Lawrence Tech.	205	320
Joseph Keefe Tech. High	208	160
Madison Park Tech. High	202	289
Minute Man Voc. Tech. High	216	199
Montachusett Voc. Tech.	210	257
Nashoba Valley Tech. High	208	121
Norfolk County Agr.	214	115
North Shore Tech. High	210	118
Northeast Metro. Reg. Voc.	206	297
Old Colony Reg. Voc. Tech.	210	127
Pathfinder Voc. Tech.	210	134
Shawsheen Valley Voc. Tech.	209	281
So. Shore Voc. Tech. High	212	130
Southeastern Reg. Voc. Tech.	209	302
Tri County Reg. Voc. Tech.	212	215
Upper Cape Cod Tech.	206	134
Whittier Reg. Voc.	204	352
Averages	210	215

Source of data: Mass. DoE, 2000h

Vocational schools typically provide instruction for grades 9 through 12 and devote about half their instructional time to traditional academics and about half to vocational training. Some communities, including Cambridge, Quincy, Revere and Waltham, provide vocational education in the same facilities as academic programs. In contrast to academic high school programs, which mainly concern themselves with preparing students for college, vocational programs train students for specific occupations, and their faculty tend to rate themselves according to graduates' success in finding satisfactory employment in those occupations.

Vocational schools, represented through the Massachusetts Association of Vocational

School Administrators, have presented the Massachusetts legislature with a bill to decouple those schools from MCAS and substitute a special examination system based on the occupational categories for which they provide training, sponsored by Sen. David Magnani of Framingham (Magnani, 2000). They are well aware, as the foregoing table shows, that their students score far below state averages on MCAS tests; but they claim that MCAS tests are directed toward a curriculum that they do not teach and cannot teach without weakening their key programs. Table A2-2 identifies some of the major sources of information for Massachusetts vocational schools available on the Internet.

Table A2-2
Massachusetts Vocational School Information Sources

Massachusetts Association of Vocational School Administrators
Bay Path Regional Vocational Technical High School, Charlton, MA
Blackstone Valley Regional Vocational Technical High School, Upton, MA
Bristol-Plymouth Regional Technical School District, Taunton, MA
Bluehills Regional Technical High School, Canton, MA
Diman Regional Vocational Technical High School, Fall River, MA
Greater Lowell Technical High School, Tyngsboro, MA
Greater New Bedford Regional Vocational Technical High School, New Bedford, MA
Lower Pioneer Valley Educational Collaborative, East Longmeadow, MA
Minuteman Science-Technology High School, Lexington, MA
Northeast Metropolitan Regional Vocational School, Wakefield, MA
Old Colony Regional Vocational Technical High School, Rochester, MA
Shawsheen Valley Technical High School. Billerica, MA
Tri-County Regional Vocational Technical School District, Franklin, MA
William J. Dean Technical High School, Holyoke, MA
Whittier Regional Vocational Technical High School, Haverhill, MA
Worcester Vocational High School, Worcester, MA
Population trends at Massachusetts Regional Vocational School Districts

Source of data: Massachusetts Association of Vocational School Administrators

Appendix 3: Metropolitan Boston MCAS Mathematics Scores

Table A3-1
1999 MCAS Grade 10 Math Test Scores for Boston-area Schools

		2000		1999		1998	
City or Town	**High School**	**Adj Avg**	**Adj Num**	**Adj Avg**	**Adj Num**	**Adj Avg**	**Adj Num**
Arlington	Arlington	239	231	234	246	230	251
Belmont	Belmont	247	254	243	229	245	223
Boston	Boston High	210	111	204	132	203	168
Boston	Brighton	208	233	205	216	203	186
Boston	Charlestown	210	175	206	148	207	156
Boston	Dorchester	205	134	204	121	206	128
Boston	East Boston	210	270	205	206	204	225
Boston	Hyde Park	204	129	203	99	203	249

Boston	Jeremiah Burke	211	112	208	145	204	85
Boston	South Boston	210	177	205	159	205	182
Boston	The English High	208	200	204	203	203	230
Boston	West Roxbury	210	211	205	269	202	149
Boston Exam	Boston Latin	262	382	254	399	246	417
Boston Exam	Latin Academy	252	248	233	245	230	222
Boston Exam	O'Bryant Science	235	222	227	258	223	248
Braintree	Braintree	241	319	228	306	229	324
Brookline	Brookline	245	429	240	416	238	406
Cambridge	Rindge & Latin*	219	331	220	423	222	404
Chelsea	Chelsea	219	203	216	207	215	242
Dedham	Dedham	238	180	227	174	226	172
Everett	Everett*	218	325	221	295	216	277
Lexington	Lexington	243	407	238	368	239	362
Lynn	Classical	223	294	216	281	216	246
Lynn	English	224	285	213	247	212	257
Malden	Malden	225	305	221	286	219	265
Marblehead	Marblehead	241	169	232	145	237	187
Medford	Medford*	228	228	221	231	219	232
Melrose	Melrose	228	214	226	245	225	243
Milton	Milton	232	241	228	208	229	204
Newton	North*	248	521	239	468	242	468
Newton	South	252	309	242	285	240	282
Peabody	Veterans*	226	426	220	354	219	372
Quincy	North Quincy	234	301	227	303	224	294
Quincy	Quincy*	220	238	212	200	214	279
Revere	Revere*	224	251	218	306	216	286
Salem	Salem*	224	234	220	243	219	257
Saugus	Saugus	233	193	226	226	222	232
Somerville	Somerville*	218	328	216	338	217	307
Stoneham	Stoneham	235	171	227	169	233	171
Swampscott	Swampscott	239	195	240	161	224	181
Wakefield	Memorial	231	192	227	228	230	243
Waltham	Waltham*	228	344	220	339	221	384
Watertown	Watertown	233	164	231	159	221	156
Weymouth	Weymouth*	227	466	222	455	222	433

Winchester	Winchester	249	225	243	198	244	178
Winthrop	Winthrop*	227	120	223	118	218	162
Woburn	Woburn	237	276	226	285	226	240

Source of data: see text, Appendix 3

Schools marked with asterisks (*) in Table A3-1 are those providing vocational education in the same facility as academic programs.

Data in Table A3-1 are tenth-grade MCAS mathematics test scores for 1998-2000, averaged by schools, and numbers of test participants per school, obtained from the Massachusetts Department of Education (Mass. DoE, 2000h) and adjusted for percentages of students enrolled in schools but not taking the test. Adjustment formulas are as follows:

$$\mathbf{N}_{\mathbf{adj}} = \mathbf{N}(1 - \mathbf{P}_{\mathbf{a}}/100)$$

$$\mathbf{S}_{\mathbf{adj}} = (100\mathbf{S} - 200\mathbf{P}_{\mathbf{a}}) / (100 - \mathbf{P}_{\mathbf{a}}),$$

where

N is the number of enrolled students

S is the average score on test, per Department of Education

$\mathbf{P}_{\mathbf{a}}$ is the percentage of "absent" students (not taking test)

$\mathbf{N}_{\mathbf{adj}}$ is the adjusted number of students (number taking test)

$\mathbf{S}_{\mathbf{adj}}$ is the adjusted average score (only students taking test)

This procedure cannot adjust correctly for students absent for some but not all test sections. The Department has not published such information for the years 1998-2000. Adjusted results are rounded to the nearest integer.

Appendix 4: Metropolitan Boston School Characteristics

Table A4-1
Data for Boston-area School Characteristics

City or Town	High School	Pop./ Grade	% African Amer.	% Asian or Pac. Isl.	% Hispanic/ Latino	% Lim. Eng. Pr.	% Free Lunch	9,10-11,12 Reduct.	Per Cap. Income
Arlington	Arlington	259	6.7	4.3	2.7	0.7	8.1	0.2	21.4
Belmont	Belmont	241	2.8	5.6	0.2	0.4	2.5	10.4	26.8

Boston	Boston High	183	56.5	3.3	26.5	14.4	64.3	4.0	15.6
Boston	Brighton	267	49.5	7.9	35.1	32.0	62.1	29.4	15.6
Boston	Charlestown	288	36.1	22.5	27.3	36.9	59.9	14.3	15.6
Boston	Dorchester	249	66.7	2.7	25.7	29.7	51.4	26.2	15.6
Boston	East Boston	317	24.0	6.7	40.3	29.8	62.9	47.0	15.6
Boston	Hyde Park	220	72.0	1.6	18.4	23.5	53.6	15.7	15.6
Boston	Jeremiah Burke	176	86.0	3.6	26.1	32.9	56.7	20.0	15.6
Boston	South Boston	267	46.7	18.4	16.6	26.1	44.6	24.3	15.6
Boston	The English High	332	50.1	1.9	36.9	45.5	62.0	24.2	15.6
Boston	West Roxbury	328	64.7	2.4	21.1	28.7	64.3	18.8	15.6
Boston Exam	Boston Latin	398	18.5	22.7	8.4	0.6	29.3	16.2	15.6
Boston Exam	Latin Academy	255	27.7	20.2	7.7	2.8	40.0	28.3	15.6
Boston Exam	O'Bryant Science	252	45.1	31.8	13.4	18.3	48.7	23.0	15.6
Braintree	Braintree	337	3.6	3.0	1.0	0.4	4.2	12.0	18.6
Brookline	Brookline	428	12.5	13.7	3.9	2.6	15.2	8.7	29.0
Cambridge	Rindge & Latin*	478	37.4	7.6	14.2	7.7	14.8	7.7	19.9
Chelsea	Chelsea	281	8.5	10.9	63.7	15.1	57.8	39.1	11.6
Dedham	Dedham	179	1.6	2.2	4.1	1.7	3.2	22.6	19.0
Everett	Everett*	336	8.4	3.6	11.2	5.2	23.0	21.1	14.2
Lexington	Lexington	386	7.3	15.0	0.6	1.1	3.9	12.3	30.7
Lynn	Classical	282	17.5	24.0	14.2	14.6	38.8	26.5	13.0
Lynn	English	327	11.3	14.2	23.0	20.1	23.2	20.5	13.0
Malden	Malden	346	13.0	21.0	6.8	10.5	17.6	15.6	15.8
Marblehead	Marblehead	178	3.9	0.7	0.6	1.1	3.5	-3.7	30.6
Medford	Medford*	266	11.2	3.9	1.3	2.9	6.0	-1.7	16.9
Melrose	Melrose	242	4.2	1.3	1.2	0.7	5.4	-2.9	20.2
Milton	Milton	232	15.7	2.2	1.3	0.8	4.2	7.9	22.4
Newton	North*	495	5.9	7.8	1.7	2.4	3.0	8.4	28.8
Newton	South	300	3.8	11.1	1.0	1.9	1.6	8.2	28.8
Peabody	Veterans*	418	0.8	1.2	7.5	2.0	8.4	20.0	17.0

Quincy	North Quincy	325	0.6	26.5	1.2	0.0	14.4	19.5	17.4
Quincy	Quincy*	282	4.7	14.1	3.6	9.0	15.7	4.7	17.4
Revere	Revere*	339	3.6	15.4	7.9	4.6	21.6	26.4	14.7
Salem	Salem*	279	5.1	2.2	20.0	6.8	26.3	31.4	16.2
Saugus	Saugus	221	0.7	1.9	0.7	0.0	7.9	17.2	17.8
Somerville	Somerville*	467	17.1	5.9	14.3	15.4	52.8	24.6	15.2
Stoneham	Stoneham	191	0.5	1.4	1.2	0.0	7.2	3.3	18.2
Swampscott	Swampscott	184	1.5	0.3	1.6	1.4	2.8	4.0	25.6
Wakefield	Memorial	240	1.9	1.5	0.4	0.1	5.8	5.7	19.0
Waltham	Waltham*	381	9.0	7.1	15.1	3.4	15.9	-2.5	16.8
Watertown	Watertown	184	3.7	3.4	4.2	3.1	14.2	6.1	20.4
Weymouth	Weymouth*	464	2.1	1.8	1.0	0.1	6.3	16.8	18.4
Winchester	Winchester	209	1.6	4.0	0.5	0.0	2.3	16.9	30.6
Winthrop	Winthrop*	140	2.3	2.1	2.7	1.6	4.1	-12.9	17.9
Woburn	Woburn	284	2.0	2.1	3.3	1.6	7.7	18.7	18.2

Source of data: see text, Appendix 4

Data in Table A4-1, except per-capita income, are gathered from the reports of public schools to the Massachusetts Department of Education for the school year ended June 30, 1999, and published by the Department in its school profiles (Mass. DoE, 2000f). Data for per-capita community income in $1,000s are from the Massachusetts Department of Revenue (Mass. DoR, 1999), based on the 1990 US Census of Population and Housing. Schools marked with asterisks (*) are those providing vocational education in the same facility as academic programs. All numbers except average student population per grade and per-capita income are percents. Grade size reduction has been calculated as percent decrease in the grades 11+12 school population as compared with grades 9+10. School district reports (Mass DoE, 1999a) include the following: (Note 38)

Foundation Enrollment and Student Attendance Reports, reporting foundation enrollment and student attendance as of October 1, submitted to the Department by each district no later than December 15 during each school year.

Individual School Reports and School System Summary Reports, reporting on student enrollments and classifications as of October 1, including gender, "racial" or "ethnic," limited English proficiency and low-income status, submitted to the Department by each district no later than January 31 during each school year.

School Attending Children Report, counting all school-age residents of a district classified by grade and type of school attended, as of January 1, submitted to the Department by each district no later than April 30 during each school year.

Year-End School Indicator Reports, submitted to the Department by each district no later than the August 15 following each school year.

End-of-Year Pupil and Financial Reports (EOYR), submitted to the Department by each district no later than the September 30 following each school year. An EOYR currently consists of the following schedules:

1. Revenue and Expenditure Summary
2. Assessments Received From Member Towns or Cities or Regional School Districts
3. Instructional Services By Grade Level
4. Special Education Expenditures by Prototype
7. Pupil Transportation Reimbursement
8. Professional Development
11. Pupil Membership Summary
13. Staff Data By Major Program Area Instructional Programs
16. Pupils-Attendance Data
19. Annual School Budget

As of July 1, 2000, the Department requires school districts to compile data on instructional costs at the school building level. New financial reporting will use a revised and uniform chart of accounts in the EOYR due September 30, 2002. However, as previously noted, current school profiles do not include expenditures for schools. The Department has a long cycle for publication of data. Final summary data on per-pupil expenditures for the school year ending June 30, 1999, were released January 20, 2001.

Acknowledgements

The author acknowledges with thanks the comments of several reviewers. I particularly appreciate Anne Wheelock's keen observations about some data interpretations. Any errors or omissions remain, as always, the responsibility of the author.

Notes

[1]The terms "racial" and "ethnic" will be used here in the senses of state and federal regulations, although they describe populations not always closely related by genetic or cultural backgrounds.

[2]For the fall of 1997, 85 percent of Massachusetts elementary and secondary school students were in public schools and 15 percent in private schools of all types. See National Center, 2001, Tables 39 and 64.

[3]On the motion of Chairman John Silber, Mass. DoE, 1998a.

[4]Massachusetts General Laws, Chapter 69 (Powers), Sections 1B and 1D. The Department of Education publishes "curriculum frameworks" with only the legal status of guidelines and recommendations. See Mass. DoE, 2000j.

[5]Massachusetts General Laws, Chapter 69 (Powers), Section 1D, paragraph (i).

[6]See Mass. DoE, 2000i, Chapter 12, for procedures to equate scores across test series. This document does not describe any procedures for relating test scores to education content as delivered in Massachusetts public schools.

[7]Model and MANOVA evaluations performed with *Statistica,* Version 5.5, StatSoft, Tulsa, OK.

[8]Compare "promoting power" or "holding power" in Balfanz and Legters, 2001. For issues concerning availability of accurate data on public school dropout rates, see Kaufman, 2001.

[9]Other school characteristics also deserve consideration. Boston may have enrollment patterns for high-school mathematics classes differing from suburban norms. Ninth-grade students classified as "Asian" have been found more than twice as likely as average students to be enrolled in traditional "high school mathematics" courses, while students classified as "Hispanic" were found less likely than average to be taking such courses. (Anne Wheelock, personal communication, January, 2001).

[10]It is possible that "limited English proficiency" acts as a proxy for other, potentially stronger factors, much as certain "racial" or "ethnic" percentages appear to act as proxies for income factors in these data.

[11]Other studies have reported significant "racial" or "ethnic" test score differences for individuals after adjusting for incomes. For one survey and an evaluation, see Hedges and Nowell, 1998, pp. 149-181. However, few studies have considered school ratings based on scholastic achievement tests given in the mid-teen years and have used commensurate data for a geographical cluster of communities that were each fairly homogeneous.

[12]See Mass. DoE, 1999b. For the first published ratings, see Mass. DoE, 2001. For potential penalties, see Mass. DoE, 2000b.

[13]See "Relationships between student MCAS performance and courses taken," in Brian Gong, *Relationships Between Student Performance on the MCAS and Other Tests--Collaborating District A,* National Center for the Improvement of Educational Assessment, Inc., March, 1999, pp. 39-48, available at http://www.nciea.org/uploadCFA/MA%20report.pdf.

[14]See Cremin, 1970, Kaestle, 1983, Tyack, 1974, and Tyack and Cuban, 1995.

[15]Referred to herein as the "Business Alliance." Not to be confused with the Massachusetts Global Business Alliance, Business Education Alliance, Regional Education and Business Alliance, or any of several other organizations with similar names.

[16]See Mass. Business, 1991. Principal author of the sections dealing with school finance was economist Edward Moskovitch, a former Massachusetts chief budget director and executive director of the Massachusetts Municipal Association. The title of this document should not be confused with the Every Child a Winner (ECAW) school programs and educational games produced by Martha F. Owens and Susan B. Rockett, beginning in 1974, and now distributed by Educational Excellence, Inc., of Ocilla, GA;

there appears to be no connection; see http://www.ed.gov/pubs/EPTW/ep tw9/eptw9b.html.

17"Students in the graduating class of 2003 shall meet or exceed the...score of 220 on both the English Language Arts and the Mathematics MCAS grade 10 tests...." See Mass. DoE, 1999f.

18See Mass. DoE, 1999c, Chapter 8, Scoring, and Chapter 10, Scaling.

19Mass. DoE, 2000a. Also see Mass DoE, 1998c, and Mass. DoE, 1999c.

20See Antonucci, 1995, for the following: "As students are better matched to their institutions," [says Abigail Thernstrom], "as they cascade to places where they are prepared to the average level, the graduation rates should go up for minorities."

21For most of the available information, see Mass DoE, 1999c.

22Massachusetts General Laws, Chapter 69 (Powers), Section 1I (second paragraph). Also see Allen, 1999.

23Information on Mass Insight is available at http://www.massinsight.com/meri/index. html.

24For example, a "Starting Now" brochure for parents. See Driscoll, 1998a, and Driscoll, 1999b. The versions distributed include: Fall 1998, Spring 1999, Manufacturing and Processing Industry Spotlight, Fall 1999, Spring 2000, Fall 2000 and Spring 2001--all available from the Massachusetts Department of Education at http://www.doe.mass.edu/mcas/p ubs/other_pub.html.

25See the following examples: "Recent survey by Mass Insight," cited in *Goals 2000 Five Year Master Plan,* Goal 5, "Create a statewide infrastructure," March, 1995; "Mass Insight report on education reform," minutes of March 22, 1996; presentations on the graduation requirement by Bill Guenther, Mass Insight Education and by Susan Kiernan and John Lozada, Mass Insight's Campaign for Higher Standards, minutes of November 23, 1999.

26"A majority of parents surveyed said they would rather pay out of their pockets to send their children to private or parochial schools than send them to Haverhill schools" (quoted in Grodsky, 1999). "Almost every school system is loaded with incompetent administrators" (John Silber, quoted in Drew and Suhler, 1998). "We should change the teacher tenure law so we can dismiss incompetent teachers" (Lamar Alexander, former Governor of Tennessee and US Secretary of Education, quoted in Patch and Wallace-Wells, 1998). "The dismissal of incompetent teachers is made almost impossible in some communities by such over-zealous delirium on the part of good people" (from Samuel P. Orth, the author of *A History of Cleveland*, in Orth, 1909).

27For passing scores, see Mass. DoE, 1999f, and Mass. DoE, 1999c, Chapter 8, Scoring, and Chapter 10, Scaling. For typical scores from 1998 and 1999, statewide, see, Mass. 1999e, Executive Summary.

[28]For passing scores, see NY DoE, 2000a. For typical scores from 1999, see NY DoE, 2000b.

[29]For passing scores, see TX DoE, 2000a. For typical scores from 2000, see TX DoE, 2000b.

[30]In some formats, New York and Texas report lower failure percentages because they exclude some students or they include students who pass after multiple attempts.

[31]See, for example, National Assessment, 1998, and similar summary tables from other years.

[32]Massachusetts General Laws, Chapter 69 (Powers), Section 1I (second paragraph).

[33]See Madaus and Clarke, 1998.

[34]See McNeil, 2000, and Heubert and Hauser, 1999.

[35]See Clarke, et al., 2000, and Haney, 2000.

[36]As required by Massachusetts General Laws, Chapter 74 (Vocational Education), Section 7, and Mass. DoE, 1997.

[37]Data from Mass. DoE, 2000h.

[38]Also see *EdTech,* Massachusetts Department of Education, at http://www.doe.mass.edu/edtech.

References

Allen, M. (1999, September 9). SouthCoast schools condemn governor's ranking plan. *New Bedford Standard-Times,* New Bedford, MA.

Antonucci, R. V. (1997a). Letter from the Commissioner of Education to school superintendents, March 18, 1997. Malden, MA: Department of Education. August, 2001, available at http://www.doe.mass.edu/mailings/1997/cm31897.html.

Antonucci, R. V. (1997b). Memorandum from the Commissioner of Education to school superintendents, May 6, 1997. Malden, MA: Department of Education. August, 2001, available at http://www.doe.mass.edu/mailings/1997/cm597ersg.html.

Antonucci, T. W. (1995). Independent Women's Forum Expose. *Women Leaders Online,* published at http://www.wlo.org/iwf/iwf2.html.

Avenoso, K. (1996, December 18). Colleagues say Silber hurting ed reform. *Boston*

Globe, p. B1.

Balfanz, R., & Legters, N. (2001). How many central city high schools have a severe dropout problem, where are they located, and who attends them? Paper presented at the Conference on Dropout Research, Harvard University, Cambridge, MA. January, 2001, draft available at http://www.law.harvard.edu/civilrights/publications/dropout/balfanz.html.

Battenfeld, J., & Pressley, D. S. (1999, March 2). Cellucci eyes replacing Silber as education czar. *Boston Herald.*

Berliner, D. C., & Biddle, B. J. (1995). *The Manufactured Crisis: Myths, Fraud, and the Attack on America's Public Schools.* Reading, MA: Addison-Wesley.

Berliner, D. C., & Biddle, B. J. (1996). Making mountains out of molehills. *Education Policy Analysis Archives 4*(3). February, 1996, available at http://epaa.asu.edu/epaa/v4n3.

Bevington, P. R., & Robinson, D. K. (1992). *Data Reduction and Error Analysis for the Physical Sciences* (2nd Ed.). New York, NY: McGraw-Hill.

Bolon, C. (2000). School-based standard testing. *Education Policy Analysis Archives 8*(23). May, 2000, available at http://epaa.asu.edu/epaa/v8n23. *Colloquia Manilana 8,* 90-145. Quezon City, Philippines.

Clarke, M., Haney, W., & Madaus, G. (2000). High stakes testing and high school completion. *National Board on Educational Testing and Public Policy Monographs 1*(3). January, 2000, available at http://nbetpp.bc.edu/statements/V1N3.pdf.

Cremin, L. A. (1970). *American Education; The Colonial Experience, 1607-1783.* New York, NY: Harper & Row.

Daley, B. (1997, October 9). Councilor questions exam process at Latin; Yancey also seeks admissions data. *Boston Globe,* p. C19.

Daley, B. (1998, May 29). Admission policy at city exam schools is upheld. *Boston Globe,* p. B1.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives 8*(1). January, 2000, available at http://epaa.asu.edu/epaa/v8n1.

Drew, D. P., & Suhler, J. N. (1998, November 15). Strong steps urged for DISD. *Dallas Morning News*.

Driscoll, D. P. (1998a). Memorandum from the Commissioner of Education to school superintendents, October 19, 1998. Malden, MA: Department of Education. August, 2001, available at http://www.doe.mass.edu/mcas/1098stnow.html.

Driscoll, D. P. (1998b). Letter from the Commissioner of Education to school superintendents, November 16, 1998. Malden, MA: Department of Education. August, 2001, available at http://www.doe.mass.edu/mailings/1998/111698.html.

Education Reform Act of 1993. Massachusetts General Court Acts of 1993, C. 71. July, 1997, available at http://www.doe.mass.edu/edreform/era1993.html.

Estrin, R (1999, March 4). Cellucci accepts Silber's resignation, breaking impass. *New Bedford Standard-Times,* New Bedford, MA.

Ferguson, R. F. (1998). Can schools narrow the black-white test score gap? In C. Jencks & M. Phillips (Eds.), *The Black-White Test Score Gap* (pp. 318-374). Washington, DC: Brookings Institution.

French, D. (1998). The state's role in shaping a progressive vision of public education. *Phi Delta Kappan 80*(3), pp. 185-194.

Future of Mass. ed. reform unclear (1996, Spring). *FairTest Examiner.* Cambridge, MA: National Center for Fair and Open Testing.

Gaudet, R. D. (2001). *Effective School Districts in Massachusetts* (3rd Ed.). Boston, MA: Donahue Institute, University of Massachusetts.

Gipps, C., & Murphy, P. (1994). *A Fair Test? Assessment, Achievement and Equity.* Philadelphia, PA: Open University Press.

Gong, B. (1999). *Relationships Between Student Performance on the MCAS and Other Tests--Collaborating District A.* Dover, NH: National Center for the Improvement of Educational Assessment, Inc. March, 1999, available at http://www.nciea.org/uploadCFA/MA%20report.pdf.

Grissmer, D. W., Flanagan, A., Kawata, J., & Williamson, S. (2000). *Improving Student Achievement: What NAEP State Test Scores Tell Us.* Santa Monica, CA: Education Division, RAND Corp. July, 2000, available at http://www.rand.org/publications/MR/MR924.

Grosky, J. B. (1999, March 5). School choice parents unhappy with Haverhill. *Lawrence Eagle-Tribune,* Lawrence, MA.

Haney, W. (2000). The myth of the Texas miracle in education. *Education Policy Analysis Archives 8*(41). August, 2000, available at http://epaa.asu.edu/epaa/v8n41.

Hayman, R. L., Jr. (1998). *The Smart Culture.* New York, NY: New York University Press.

Hedges, L. V., & Nowell, A. (1998). Black-white test score convergence since 1965. In C. Jencks & M. Phillips (Eds.), *The Black-White Test Score Gap* (pp. 149-181). Washington, DC: Brookings Institution.

Heubert, J. P., & Hauser, R. M., Eds. (1999). *High Stakes Testing for Tracking, Promotion and Graduation.* Washington, DC: National Academy Press.

Horn, C., Ramos, M., Blumer, I., & Madaus, G. (2000). Cut scores: Results may vary. *National Board on Educational Testing and Public Policy Monographs 1*(1). April, 2000, available at http://nbetpp.bc.edu/statements/M1N1.pdf.

Howe, P. J. (1992, December 5). Cellucci wants plan for schools on table, urges legislators to reveal details. *Boston Globe,* p. 13.

Howe, P. J. (1993a, January 27). School reform debate begins but antiabortion phrases cause stir. *Boston Globe,* p. 14.

Howe, P. J. (1993b, January 29). House in 4th day of talks on education reform bill. *Boston Globe,* p. 13.

Howe, P. J. (1993c, April 22). Education reform loses luster on the hill. *Boston Globe,* p. 57.

Jackson, D. Z. (1996, August 9). A warped board of education. *Boston Globe,* p. A19.

Kaestle, C. F. (1983). *Pillars of the Republic.* New York, NY: Hill and Wang.

Kaufman, P. (2001). The national dropout data collection system: Assessing consistency. Paper presented at the Conference on Dropout Research, Harvard University, Cambridge, MA. January, 2001, draft available at http://www.law.harvard.edu/civilrights/publications/dropout/kaufman.html.

Klein, S. P., Hamilton, L. S., McCaffrey, D. F., & Stecher, B. M. (2000). What Do Test Scores in Texas Tell Us?*Education Policy Analysis Archives 8*(49). October, 2000, available at http://epaa.asu.edu/epaa/v8n49.

Kohn, A. (2000, March 20). Poor teaching for poor students: more reasons to boycott the MCAS tests. *Boston Globe,* p. A11.

Koretz, D. (1992). What happened to test scores, and why? *Educational Measurement: Issues and Practice 11*, pp. 7-11.

Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics 114,* 497-532.

Langborgh, E. (1999). CUNY remediates itself. *Accuracy in Academia,* published at http://www.academia.org/cunyremediation.html.

Lehigh, S. (1993, March 13). Roosevelt may rouse interest in '94. *Boston Globe,* p. 14.

Leung, S. (1997, January 16). State board rejects GED exam plan. *Boston Globe,* p. B3.

Madaus, G., & Clarke, M. (1998). The adverse impact of high stakes testing on minority students: Evidence from 100 years of test data. Paper presented at the High Stakes K-12 Testing Conference, Teachers College, Columbia University, New York, NY. December, 1998, draft available at http://www.law.harvard.edu/civilrights/conferences/testing98/drafts/madaus2.html

Magnani, D. P., Havern, R. A., Resor, P. A., et al. (2000). An act recognizing vocational-technical learning in educational assessment. Massachusetts Senate Bill 282, Senate Docket 1656. December, 2000, available at http://www.state.ma.us/legis/bills/st00282.htm.

Marcus, J. (1999, October). The shocking truth about our public schools. *Boston Magazine.*

Massachusetts Business Alliance for Education (1991). *Every Child a Winner: A Proposal for a Legislative Action Plan for Systemic Reform of Massachusetts' Public Primary and Secondary Education System.* Worcester, MA: Massachusetts Business Alliance for Education.

Massachusetts Department of Education (1994). *Massachusetts Common Core of Learning,* Malden, MA: Department of Education. August, 2001, available at http://www.doe.mass.edu/edreform/commoncore/purpose.html.

Massachusetts Department of Education (1996a). High expectations for students and accountability for all. *Education Today 12*(1). Malden, MA: Department of Education.

Massachusetts Department of Education (1996b). Assessment development committees represent school, district and community interests. *Education Today 12*(1). Malden, MA: Department of Education.

Massachusetts Department of Education (1997). Vocational education. Malden, MA: Department of Education, Robert V. Antonucci, Commissioner. Massachusetts Regulation 603 CMR 4.00 et seq., July, 1997.

Massachusetts Department of Education (1998a). *Minutes of the Massachusetts Board of Education,* January 12, 1998. Malden, MA: Department of Education. Summary available at http://www.doe.mass.edu/boe/bib/bib98/bib11298.html.

Massachusetts Department of Education (1998b). *MCAS Performance Level Definitions.* Malden, MA: Department of Education, David P. Driscoll, Commissioner. September, 1998, available at http://www.doe.mass.edu/mcas/mcaspld.html.

Massachusetts Department of Education (1998c). *Background on the MCAS tests of May, 1998.* Malden, MA: Department of Education, David P. Driscoll, Commissioner. November, 1998, available at http://www.doe.mass.edu/mcas/1998bg/Overview_TableContents.html.

Massachusetts Department of Education (1999a). Accounting and reporting: school districts. Malden, MA: Department of Education, David P. Driscoll, Commissioner. Massachusetts Regulation 603 CMR 10.03, July, 1999.

Massachusetts Department of Education (1999b). *School and District Accountability System,* Malden, MA: Department of Education, David P. Driscoll, Commissioner. October, 1999, available at http://www.doe.mass.edu/ata/pdf/sdas99.pdf.

Massachusetts Department of Education (1999c). *MCAS 1998 Technical Report,* Malden, MA: Department of Education, David P. Driscoll, Commissioner. September, 1999, available at http://www.doe.mass.edu/mcas/archives/99/98tech_report_full.pdf.

Massachusetts Department of Education (1999d). *Massachusetts Comprehensive Assessment System: Guide to Interpreting the 1999 MCAS Reports for Schools and Districts,* Malden, MA: Department of Education, David P. Driscoll, Commissioner.

November, 1999, available at
http://www.doe.mass.edu/mcas/99results/interp99/full_guide.pdf.

Massachusetts Department of Education (1999e). *Massachusetts Comprehensive Assessment System: Report of 1999 State Results,* Malden, MA: Department of Education, David P. Driscoll, Commissioner. November, 1999, available at http://www.doe.mass.edu/mcas/99results/99mcas/110899com.html.

Massachusetts Department of Education (1999f). Standards for competency determination. Malden, MA: Department of Education, David P. Driscoll, Commissioner. Massachusetts Regulation 603 CMR 30.03, November, 1999.

Massachusetts Department of Education (2000a). Massachusetts certificate of mastery. Malden, MA: Department of Education, David P. Driscoll, Commissioner. Massachusetts Regulation 603 CMR 31.03, March, 2000.

Massachusetts Department of Education (2000b). Underperforming schools and school districts. Malden, MA: Department of Education, David P. Driscoll, Commissioner. Massachusetts Regulation 603 CMR 2.00 et seq., May, 2000.

Massachusetts Department of Education (2000c). *Report of 1999 Massachusetts and Local School District MCAS Results by Race/Ethnicity.* Malden, MA: Department of Education, David P. Driscoll, Commissioner. May, 2000, available at http://www.doe.mass.edu/mcas/race_report99/default.html.

Massachusetts Department of Education (2000d). *Massachusetts Comprehensive Assessment System Requirements for the Participation of Students with Limited English Proficiency.* Malden, MA: Department of Education, David P. Driscoll, Commissioner. June, 2000, available at http://www.doe.mass.edu/mcas/2000docs/pdf/lep.pdf.

Massachusetts Department of Education (2000e). *Final FY01 Chapter 70 Aid and Spending Requirements.* Malden, MA: Department of Education, David P. Driscoll, Commissioner. August, 2000, available at
http://www.doe.mass.edu/mcas/2000docs/pdf/lep.pdf.

Massachusetts Department of Education (2000f). *School and District Profiles.* Malden, MA: Department of Education, David P. Driscoll, Commissioner. September, 2000, available at http://profiles.doe.mass.edu.

Massachusetts Department of Education (2000g). *Release of Spring 2000 Test Items.* Malden, MA: Department of Education, David P. Driscoll, Commissioner. November, 2000, available at http://www.doe.mass.edu/mcas/pubs/test_rel.html.

Massachusetts Department of Education (2000h). *MCAS Results, 1998-2000.* Malden, MA: Department of Education, David P. Driscoll, Commissioner. November, 2000, available at http://www.doe.mass.edu/mcas/results.html.

Massachusetts Department of Education (2000i). *MCAS 1999 Technical Report.* Malden, MA: Department of Education, David P. Driscoll, Commissioner. November, 20010, available at http://www.doe.mass.edu/mcas/2000docs/pdf/99techrep.pdf.

Massachusetts Department of Education (2000j). *Massachusetts Curriculum*

Frameworks. Malden, MA: Department of Education, David P. Driscoll, Commissioner. November, 2000, available at http://www.doe.mass.edu/frameworks/current.html.

Massachusetts Department of Education (2000k). *Curriculum Framework Archive.* Malden, MA: Department of Education, David P. Driscoll, Commissioner. November, 2000, available at http://www.doe.mass.edu/frameworks/current.html.

Massachusetts Department of Education (2000l). *MCAS Alternate Assessment.* Malden, MA: Department of Education, David P. Driscoll, Commissioner. November, 2000, available at http://www.doe.mass.edu/mcas/2000docs/MCAS_alt300.htm.

Massachusetts Department of Education (2001). *School Performance Ratings by District.* Malden, MA: Department of Education, David P. Driscoll, Commissioner. January, 2001, available at http://www.doe.mass.edu/ata/ratings00/SPRPStateSumm00.pdf.

Massachusetts Department of Revenue (1999). *Measures of Property Wealth and Income.* Boston, MA: Division of Local Services, Massachusetts Department of Revenue. September, 1999, available at http://www.state.ma.us/dls/mdmstuf/Wealth.xls.

McDuffy v. Robertson, 615 N.E.2d 516 (MA 1993).

McLaughlin v. Boston School Committee, 938 F.Supp. 1001, 1018 (MA 1996).

McNeil, L. M. (2000), *Contradictions of School Reform,* New York, NY: Routledge.

Metropolitan Area Planning Council (1997). *Sustainable Development in the Metropolitan Boston Region.* Boston, MA: Metropolitan Area Planning Council, Grace S. Shepard, President. August, 1997, available at http://info.cnt.org/mi//bo_briefing.htm.

National Assessment of Educational Progress (1998). *National and State Summary Data Tables.* Washington, DC: National Assessment of Educational Progress, US Department of Education, July 1998, available at http://nces.ed.gov/nationsreportcard/site/home.asp.

National Center for Education Statistics (2001). *Digest of Education Statistics, 2000.* Washington, DC: U. S. Department of Education.

Nealon, P. (1993, April 15). Project may run some public Mass. schools. *Boston Globe,* p. 39.

New York State Education Department (2000a). *Understanding Your School Report Card 2000, Guide to Secondary School Examinations.* Albany, NY: Education Department. February, 2000, available at http://www.emsc.nysed.gov.

New York State Education Department (2000b). *New York State School Report Card Database.* Albany, NY: Education Department. March, 2000, available at http://www.emsc.nysed.gov.

Nicodemus, A. (2000, November 21). Alternative test sought for Voc-Tech. *New Bedford Standard-Times,* New Bedford, MA.

Nye, B., Hedges, L. V., & Konstantopoulos, S. (1999). The long-term effects of small classes: A five-year follow-up of the Tennessee class size experiment. *Educational*

Evaluation and Policy Analysis 21(2), 127-142.

Orfield, G., Arenson, J., Jackson, T., et al. (1997). *City-Suburban Desegregation: Parent and Student Perspectives in Metropolitan Boston.* Cambridge, MA: Harvard Civil Rights Project, Harvard University. September, 1997, draft available at http://gseweb.harvard.edu/~apsp/oldsite/septrep.html.

Orth, S. P. (1909, March). Plain facts about public schools. *Atlantic Monthly*.

Overdue school reform bill (1992, December 23). *Boston Globe,* p. 14.

Partee, G. (1997). *High School Reform and Systemic Districtwide Reform in Boston, Massachusetts.* Washington, DC: American Youth Policy Forum. April, 1997, available at http://www.aypf.org/tripreports/1997/tr040397.htm.

Patch, B., & Wallace-Wells, B. (1998, April 22). Lamar Alexander on educating America. *Dartmouth Review,* Dartmouth, NH.

Pawlak-Seaman, S. (1996, August 9). Silber as ed czar? That's really scary. *New Bedford Standard-Times,* New Bedford, MA.

Pawlak-Seaman, S. (1997, January 8). Silber: Test scores "pathetic." *New Bedford Standard-Times,* New Bedford, MA.

Peyser, J. A. (1996). Changing the monopoly structure of public education. *Dialogue 12.* Boston, MA: Pioneer Institute.

Peyser, J. A. (1998). *Human Services Financing: A New Model.* Boston, MA: Pioneer Institute.

Phillips, M., Crouse, J., & Ralph, J. (1998). Does the black-white test score gap widen after children enter school? In C. Jencks & M. Phillips (Eds.), *The Black-White Test Score Gap* (pp. 229-272). Washington, DC: Brookings Institution.

Powers and Duties of the Department of Education. Massachusetts General Laws, Chapter 69. Available at http://www.state.ma.us/legis/laws/mgl/gl-69-toc.htm.

Public Schools. Massachusetts General Laws, Chapter 71. Available at http://www.state.ma.us/legis/laws/mgl/gl-71-toc.htm.

Ribadeneira, D. (1993, January 26). Coalition raps aspects of reform package. *Boston Globe,* p. 14.

Rowe, K. J. (1999a). Assessment, performance indicators, "league tables," "value-added" measures and school effectiveness. Paper presented at the Australian Association for Research in Education 1999 Conference, Melbourne, Australia. December, 1999, available at http://www.swin.edu.au/aare/99pap/row99656.htm.

Rowe, K. J., Turner, R., & Lane, K. (1999b). The "myth" of school effectiveness: locating and estimating the magnitudes of major sources of variation in students' year 12 achievements within and between schools over five years. Paper presented at the Australian Association for Research in Education 1999 Conference, Melbourne,

Australia. December, 1999, available at
http://www.swin.edu.au/aare/99pap/row99125.htm.

Sacks, P. (1999). *Standardized Minds.* Cambridge, MA: Perseus Books.

School reform bill benefits and costs for fiscal 1994 (1992, December 23). *Boston Globe,* p. 20.

Smith, J. R., Brooks-Gunn, J., & Klebanov, P. K. (1997). Consequences of living in poverty for young children's cognitive and verbal ability and early school development. In G. J. Duncan & J. Brooks-Gunn (Eds.), *Consequences of Growing Up Poor* (pp. 132-167). New York, NY: Russell Sage Foundation.

Steinberg, J. (2000, April 13). Blue books closed, students protest state tests. *New York Times,* p. A1.

Still too early to consider graded diploma system (1999, August 9). *Springfield Union-News,* Springfield, MA.

Stotsky, S. (1999). *Losing Our Language: How Multicultural Classroom Instruction is Undermining Our Children's Ability to Read, Write and Reason.* New York, NY: Simon & Schuster.

Szechenyi, C. (1998, April 5). Steps urged to avoid test errors; Former education commissioner recommended national oversight panel. *Middlesex News,* Framingham, MA.

Taylor, J. (1993, May 2). Crusade for schools puts CEO in limelight he shuns. *Boston Globe* (Northwest Edition), p. 1.

Texas Education Agency (2000a). *Student Assessment Division Performance Standards.* Austin, TX: Education Agency. March, 2000, available at http://www.tea.tx.us.

Texas Education Agency (2000b). *2000 Statewide TAAS Summary Reports.* Austin, TX: Education Agency. March, 2000, available at http://www.tea.tx.us.

Tuerck, D.G. (2001a). *Promoting Good Schools through Wise Spending.* Boston, MA: Beacon Hill Institute, Suffolk University. January, 2001, summary available at http://www.beaconhill.org/BHIStudies/EdStudyexecsum.pdf.

Tuerck, D.G. (2001b, January 20). MCAS rating system needs to be fixed. *Boston Globe,* p. A15.

Tyack, D. B. (1974). *The One Best System.* Cambridge, MA: Harvard University Press.

Tyack, D. B., & Cuban, L. (1995). *Tinkering Toward Utopia,* Cambridge, MA: Harvard University Press.

Uriarte, M., & Chavez, L. (1999). *Latino Students and Massachusetts Public Schools.* Boston, MA: Mauricio Gaston Institute for Latino Community Development and Public Policy, University of Massachusetts. November, 1999, available at http://www.gaston.umb.edu/mcas/mcas_mp.html.

Vaishnav, A. (2000, May 26). MCAS opponents administer a pop quiz: Questions stump many passersby. *Boston Globe,* p. B4.

Vennochi, J. (1993, March 12). Dangerous liaisons. *Boston Globe,* p. 5.

Vocational Education. Massachusetts General Laws, Chapter 74. Available at http://www.state.ma.us/legis/laws/mgl/gl-74-toc.htm.

Walser, N. (1997). Financing fairness. *Commonwealth 2*(2-3), Spring. Boston, MA: MassINC.

Wheelock, A. (1999). High-stakes test-based accountability policies: problems and pitfalls. Cambridge, MA: National Center for Fair and Open Testing. December, 1999, available at http://fairtest.org/care/High-Stakes_Test-Based_Accountability.html.

Why standards and tests? (1999). *Mass Insight Education's Organization and Programs.* Boston, MA: Mass Insight Education and Research Institute. December, 1999, available at http://www.massinsight.com/meri/e_why.html.

Wong, D. S. (1996, August 10). Weld picks education panel nominees. *Boston Globe,* p. B5.

Wood, J. (1999). *An Early Examination of the Massachusetts Charter School Initiative.* Boston, MA: Donahue Institute, University of Massachusetts. December, 1999, available at http://www.massedreformreview.org/research/charters.html.

About the Author

Craig Bolon
Planwright Systems Corp
P.O. Box 370
Brookline, MA
02446-0003 USA

Email: cbolon@planwright.com
Phone: 617-277-4197

Craig Bolon is President of Planwright Systems Corp., a software development firm located in Brookline, Massachusetts, USA. After several years in high energy physics research and then in biomedical instrument development at M.I.T., he has been an industrial software developer for the past twenty years. He is author of the textbook *Mastering C* (Sybex, 1986) and of several technical publications and patents. He is an elected Town Meeting Member and has served as member and Chair of the Finance Committee in Brookline, Massachusetts.

The World Wide Web address for the *Education Policy Analysis Archives* is **epaa.asu.edu**

General questions about appropriateness of topics or particular articles may be addressed to the Editor, Gene V Glass, glass@asu.edu or reach him at College of Education, Arizona State University, Tempe, AZ 85287-0211. (602-965-9644). The Commentary Editor is Casey D. Cobb: casey.cobb@unh.edu .