

# Chapter Eight

## Seven Hills Charter School

### Worcester, Massachusetts

#### 8.1 Descriptive Summary of School

Seven Hills Charter School is an independent charter school established in 1996 to serve grades K-8. The school opened with K-6, and in 1997 grades 7 and 8 were added. The mission of the Seven Hills

Charter School is to prepare a diverse cross section of Worcester children for success with a high quality education at prevailing public school costs.

During the 1998/99 school year, 662 students were enrolled in the school. Of these, 19.3 percent were African American, 0.6 percent were of Asian descent, 25.1 percent were Hispanic, 1.4 percent were Native American, and 53.6 percent were white. The total proportion of minorities is similar to the district of Worcester, although the district has fewer African-American students and more students of Hispanic and Asian descent.

Edison's second annual report of student achievement (Edison, 1999), noted that there were 61 instructional staff members, with 10.9 pupils per staff member. The district's school profiles, as well as the Seven Hills' profile in the state's annual report (Massachusetts Department of Education [MDOE], 1999) on charter schools, indicated that the number of instructional staff was 36, with an average of 19.8 pupils per staff member. The latter figure is slightly higher than district and state averages, which are 16.6 and 18.1 pupils per staff member, respectively.

According to data obtained from the school district profiles (MDOE, 1999), Seven Hills Charter School has a slightly lower proportion of students with special educational needs (15.5 percent in 1998/99) than the local school district (17.8 percent) and the state average (16.6 percent). Edison (1999) reports that student mobility in 1997-98 was 6.7 percent. The school has slightly fewer students qualifying for free or reduced lunch (46.4 percent) than the district (50.2 percent) (MDOE, 1999), but more than the state as a whole (25.8 percent).

In terms of students with limited English proficiency, it was difficult to determine a clear figure for Seven Hills Charter School. In the 1997/98 school year, the school had 11.5 percent of its students with limited English proficiency (LEP) while in the following year, the school reportedly had 0.5 percent (MDOE, 1999). Meanwhile, the district reported that 7 percent of its students are LEP whereas the state reported 4.7 percent (MDOE, 1999). Edison (1999) reported that 10.4 percent of its students were categorized as LEP during the 1997/98 school year. Since the district has a higher proportion of Hispanic students, the proportion of Edison's students labeled as LEP most likely does not exceed the district average, although there are incomplete and conflicting figures.

The school profile in the state charter school report (MDOE, 1998) indicated that the size of the waiting list was reported to be 418 students. The report on charter schools included two figures for the total number of days of instruction (190 and 200). The estimated total hours of structured learning time was 1,239 hours per year.

The expenditures per pupil were nearly identical for Seven Hills (\$6,446) and the district (\$6,433) in 1997/98 (MDOE, 1999). The state average was slightly lower at \$6,361 per pupil in that same year. Likewise, average attendance rates were nearly identical between Seven Hills (94.6 percent), the district (93.5 percent), and the state average (93.9 percent) (MDOE, 1999).

From the school's descriptive data that we reviewed, we can see that the size and composition of the school have not altered much since it opened in 1996. In later sections of this chapter we will be comparing the performance of students in this school with the district and the state. Therefore, it is important to consider the differences in key background characteristics between the school, district, and the state as a whole. As an urban district, both the Edison school and the district have a higher proportion of minorities and students qualifying for free or reduced lunches than the state average.

## 8.2 Past Studies and Evaluations and Available Data for Analysis

Tests administered at Seven Hills include the Massachusetts Comprehensive Assessment System (MCAS) in Spring 1998 and Spring 1999 for grades 3,4,8,10; the Stanford Achievement Test Series-9<sup>th</sup> Edition (SAT-9) in Spring 1997 for grades 3-4, in Spring 1998 for grades 3-5; the Metropolitan Achievement Test-7<sup>th</sup> Edition (MAT-7) in Spring 1997 for grades 5-7; and the Iowa Test for Basic Skills (ITBS) in Spring 1997 for grade 3 and in Spring 1998 for grade 3.

A comparison of results on the ITBS, conducted by the Worcester School Department (1997) indicated that third grade students in a district school performed considerably better on the ITBS reading subtest than did students at Seven Hills. The same study compares scores on the ITBS reading subtest among charter schools in the state. The percentile ranking of Seven Hills was second from the bottom (9<sup>th</sup> out of 10) with a percentile rank of 40, compared with the highest ranking charter school with a percentile rank of 78.

Edison provided us with an SPSS data set containing SAT-9 test data for two years (1998-1999) and three years of student achievement data on the MAT-7 (1997-1999). Similar to all Massachusetts public schools, this school is also required to take part in the Massachusetts Comprehensive Assessment System (MCAS). We secured test result data for only two years, 1997-98 and 1998-99. This matches the CRT results reported by Edison in its 2000 annual school performance report.

### 8.3 Longitudinal Analysis of Individual Student Data

Individual student achievement data on SAT-9 and MAT-7 test results were available for our evaluation. The MAT-7 data set represented a longer time span for our longitudinal analysis (i.e., three academic years rather than two); thus, we place greater confidence in these analyses. We have summarized both the MAT-7 and SAT-9 results below.

#### MAT-7

The MAT-7 scores were provided in four separate scales: grade equivalent scores (GE), standard scores (SS), national percentile rank scores (PR), and normal curve equivalent scores (NCE) for language, math, and reading. Parallel analyses are reported for each subtest and scale.

Table 8:1 presents the various sample size breakdowns by subject tested, grade, and group for 324 students covering the 1997-1999 academic years. It was possible to trace 2 different cohorts. Cohort A included 44 students over 3 consecutive academic years (i.e., grade 5 in 1996/97, grade 6 in 1997/98, and grade 7 in 1998/99). Cohort B was much smaller in size and contained only 24 students across 3 academic years (grade 6 in 1996/97, grade 7 in 1997/98, and grade 8 in 1998/99). We debated whether or not we should report the results from Cohort B because of its small size, but have opted to do so in order to include—at this level of the analysis—a larger picture of the results. Nonetheless, the results from Cohort B should be interpreted more cautiously.

Similar to many of the other Edison schools in this study, there appears to be either a high rate of attrition or substantial inconsistencies in who is taking the tests at Seven Hills. In regard to Cohort A, 69 possible students could be followed in 1997. By 1998 about 87 percent remained (60/69 pairwise analysis), and by 1999 only about 61 percent of the students remained (longitudinal trend analysis). This represents a 39 percent drop in students over 3 years in the cohort of students that progressed from grade 5 to grade 7 between the 1997/98 school year and the 1998/99 school year. The pattern for Cohort B was similar, with 40 students starting in sixth grade in 1997, dropping to 28 seventh graders (70 percent) and to 24 eighth graders in 1999 (60 percent), or a 40 percent drop in the three years.

Table 8:1 Table of Sample Sizes for Individual Student Data on the MAT-7 by Grade and Year

<b>Grade 5</b>	1997			1998			1999		
	Language	Math	Reading	Language	Math	Reading	Language	Math	Reading
GE	69	69	69						
SS	69	69	69						
PR	69	69	69						
NCE	69	69	69						
<b>Grade 6</b>	1997			1998			1999		
	Language	Math	Reading	Language	Math	Reading	Language	Math	Reading
GE	40	40	41	78	78	78	55	55	55
SS	40	40	41	78	78	78	55	55	55
PR	40	40	41	78	75	77	55	55	55
NCE	40	40	41	78	75	77	55	55	55
<b>Grade 7</b>	1997			1998			1999		
	Language	Math	Reading	Language	Math	Reading	Language	Math	Reading
GE	46	46	46	78	78	78	81	81	81
SS	46	46	46	78	78	78	81	81	81
PR	46	46	46	74	77	74	81	81	81
NCE	46	46	46	74	77	74	81	81	81
<b>Grade 8</b>	1997			1998			1999		
	Language	Math	Reading	Language	Math	Reading	Language	Math	Reading
GE				51	51	51	78	78	78
SS				51	51	51	78	78	78
PR				51	50	51	78	77	78
NCE				51	50	51	78	77	78

Sample sizes for 1 year gains (grade 5 to grade 6)

	1997/98		
	Language	Math	Reading
GE	58	55	57
SS	58	55	57
PR	58	55	57
NCE	58	55	57

Sample sizes for 1 year gains (grade 6 to grade 7)

	1997/98			1998/99		
	Language	Math	Reading	Language	Math	Reading
GE	28	29	30	62	59	61
SS	28	29	30	62	59	61
PR	28	29	30	62	59	61
NCE	28	29	30	62	59	61

	Sample sizes for 1 year gains (grade 7 to grade 8)			Sample sizes for 1 year gains (grade 7 to grade 8)		
	1997/98			1998/99		
	Language	Math	Reading	Language	Math	Reading
GE	37	37	37	61	62	61
SS	37	37	37	61	62	61
PR	37	37	37	61	62	61
NCE	37	37	37	61	62	61

**Cohort A** Sample Sizes for Longitudinal Analysis (Grades 5 to 6 to 7) 1997/98/99

	Language	Math	Reading
GE	44	41	43
SS	44	41	43
PR	44	41	43
NCE	44	41	43

**Cohort B** Sample Sizes for Longitudinal Analysis (Grades 6 to 7 to 8) 1997/98/99

	Language	Math	Reading
GE	24	24	24
SS	24	24	24
PR	24	24	24
NCE	24	24	24

### Longitudinal analysis findings

Repeated measures ANOVAs were examined for longitudinal trends over a three-year period for Seven Hills. Outcome data represented individual student data on the MAT-7 subtests: language, mathematics “concepts/problem solving,” and reading comprehension. Parallel analyses are reported for all types of scores reported by Edison, e.g., GE, SS, PR, and the NCE scores. In all models, the assumption of sphericity was evaluated, and if found to be violated, the Huynh-Feldt adjusted p-values are reported. If the overall linear model was found to be statistically significant, unadjusted (alpha) pair-wise comparisons were examined to identify where a difference in the means might be located.

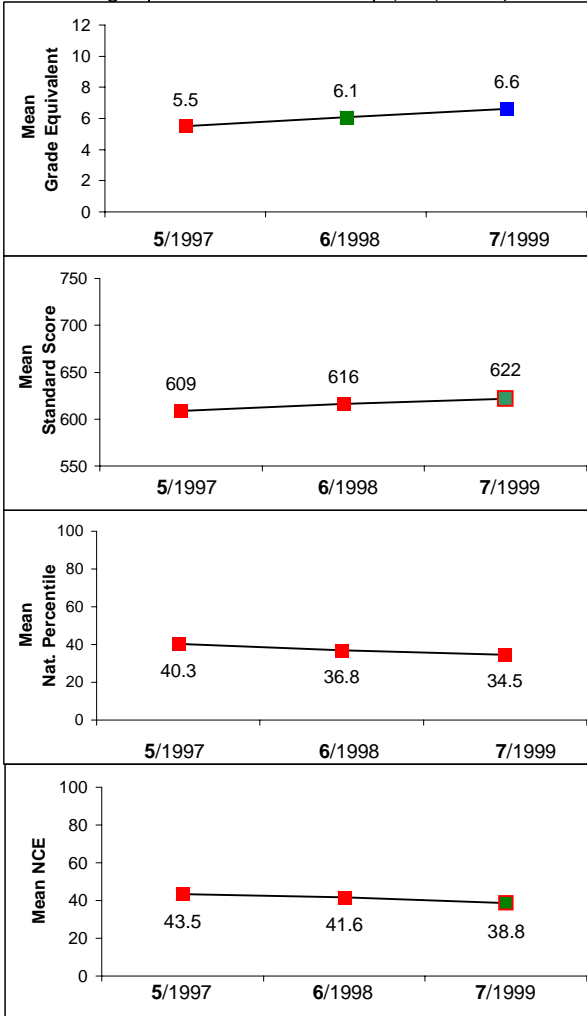
The tables and charts in Exhibits 8:1 and 8:2 illustrate the results in terms of grade level equivalents, standard scores, national percentiles, and normal curve equivalents for Cohorts A and B, respectively. The results are grouped by subject area tests: language, math/problem solving, and reading comprehension.

The students in Cohort A who had been enrolled at this school since it opened cannot be characterized as high performing students. On the whole they performed at a level that would be expected for students in their grade, although in comparison with national norms, the students were typically in the 35<sup>th</sup> to 43<sup>rd</sup> percentile. What is of particular importance for our analysis is not where they were at one point in time; rather, we are interested in the rate of learning or in the relative size of the gains they made each year.

# Exhibit 8:1 Seven Hills Charter School, Cohort A: Analysis of Individual Student Results on MAT-7

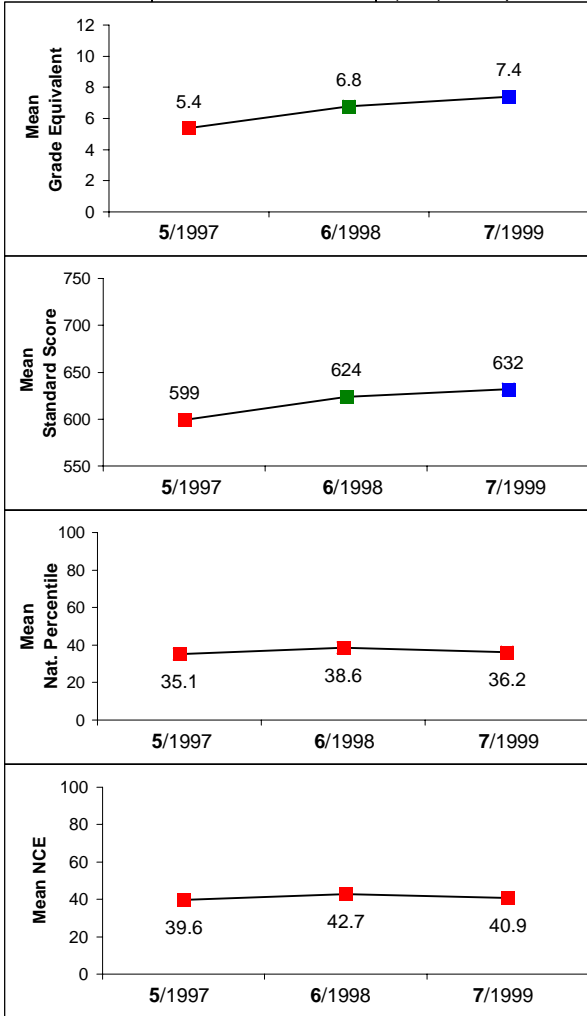
## COHORT A LANGUAGE

n=44	1997	1998	1999	
Grade	5th	6th	7th	
GE Lang	5.5	6.1	6.6	$F(3,86)=9.14, p=.0003$
SS Lang	609	616	622	$F(3,86)=8.31, p=.0005$
NP Lang	40.3	36.8	34.5	$F(3,86)=2.62, p=.0788$
NCE Lang	43.5	41.6	38.8	$F(3,86)=3.15, p=.0506$



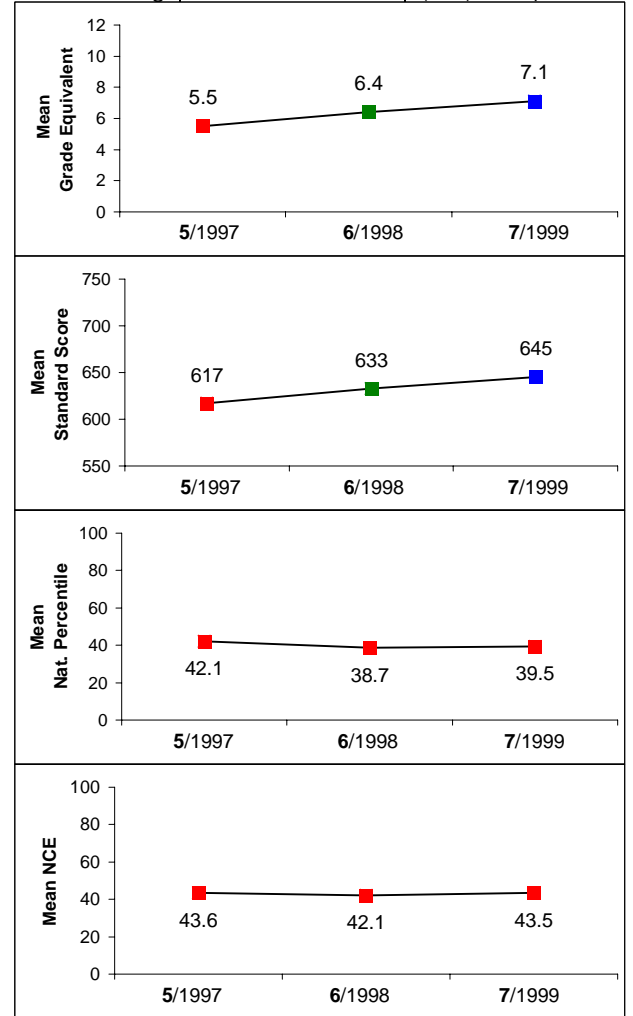
## MATH

n=41	1997	1998	1999	
Grade	5th	6th	7th	
GE Math	5.4	6.8	7.4	$F(2,80)=31.13, p<.0001$
SS Math	599	624	632	$F(2,80)=60.22, p<.0001$
NP Math	35.1	38.6	36.2	$F(2,80)=0.85, p=.4311$
NCE Math	39.6	42.7	40.9	$F(2,80)=1.13, p=.3289$



## READING

n=43	1997	1998	1999	
Grade	5th	6th	7th	
GE Reading	5.5	6.4	7.1	$F(2,84)=31.65, p<.0001$
SS Reading	617	633	645	$F(2,84)=41.07, p<.0001$
NP Reading	42.1	38.7	39.5	$F(2,84)=1.33, p=.2690$
NCE Reading	43.6	42.1	43.5	$F(2,84)=0.55, p=.5776$



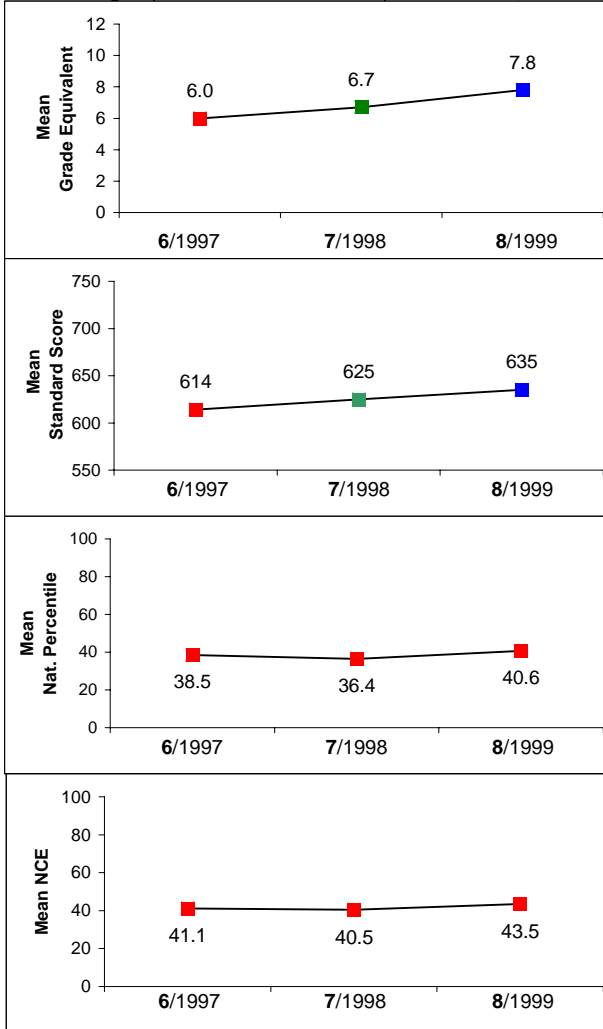
A color change represents a statistically significant change in the means. A trend with two color changes represents differences among all three means. A dual colored charting point (red and green) represents a statistically significant difference between one mean but not the other. The reader is encouraged to examine the individual table of means for these cases.

## Exhibit 8:2 Seven Hills Charter School, Cohort B: Analysis of Individual Student Results on MAT-7

### COHORT B

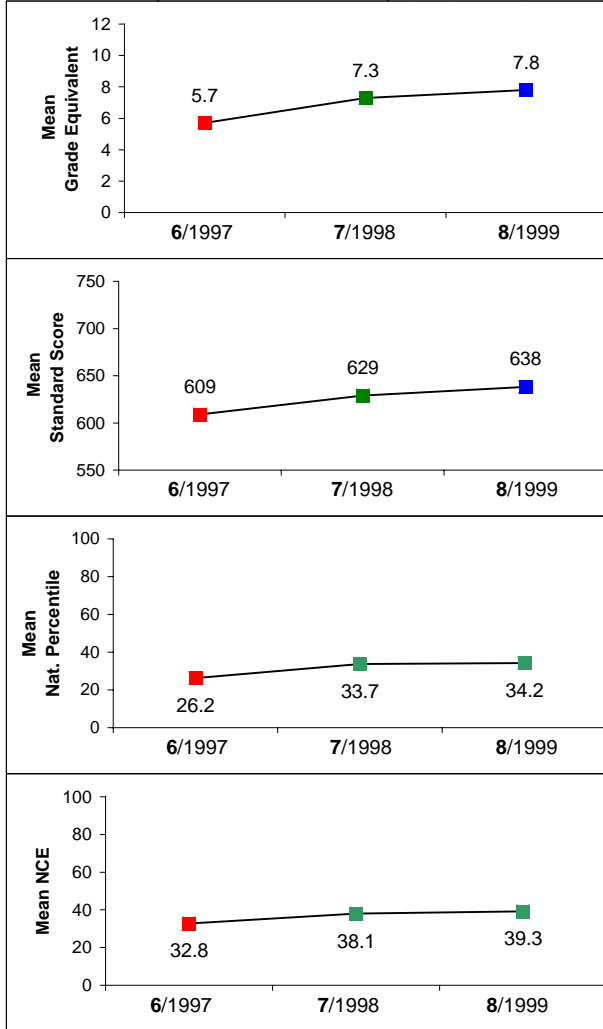
#### LANGUAGE

n=24	1997	1998	1999	
Grade	6th	7th	8th	
GE Lang	6.0	6.7	7.8	F(2,46)=12.82, p<.0001
SS Lang	614	625	635	F(2,46)=13.69, p<.0001
NP Lang	38.5	36.4	40.6	F(2,46)=0.92, p=.4056
NCE Lang	41.1	40.5	43.5	F(2,46)=0.99, p=.3777



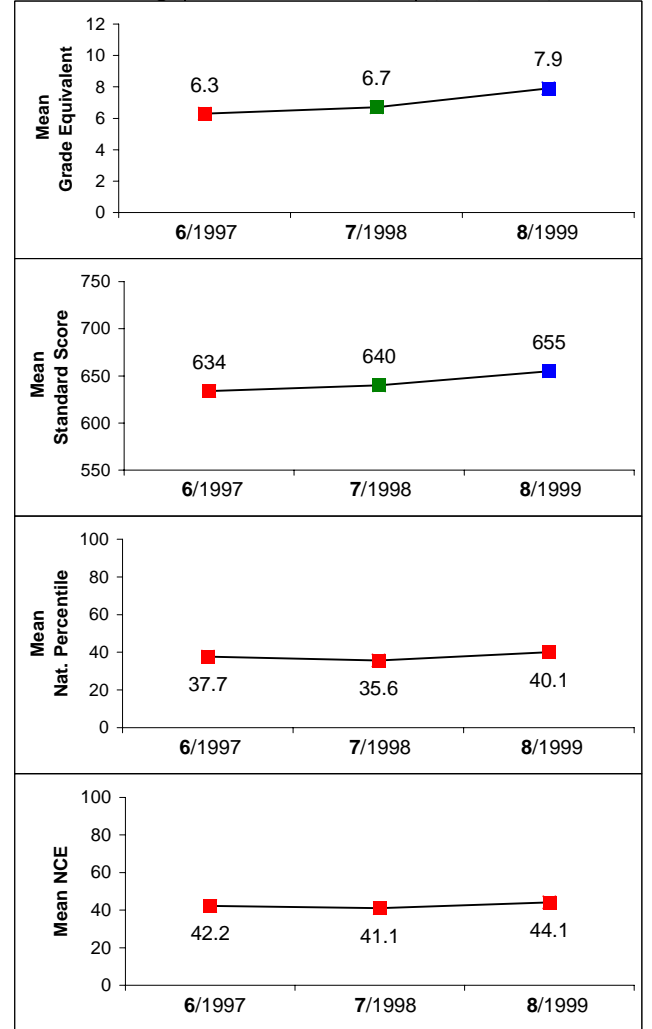
#### MATH

n=24	1997	1998	1999	
Grade	6th	7th	8th	
GE Math	5.7	7.3	7.8	F(2,46)=24.30, p<.0001
SS Math	609	629	638	F(2,46)=33.06, p<.0001
NP Math	26.2	33.7	34.2	F(2,46)=5.34, p=.0082
NCE Math	32.8	38.1	39.3	F(2,46)=4.48, p=.0167



#### READING

n=24	1997	1998	1999	
Grade	6th	7th	8th	
GE Reading	6.3	6.7	7.9	F(2,46)=21.02, p<.0001
SS Reading	634	640	655	F(2,46)=18.70, p<.0001
NP Reading	37.7	35.6	40.1	F(2,46)=1.97, p=.1515
NCE Reading	42.2	41.1	44.1	F(2,46)=1.18, p=.3174



A color change represents a statistically significant change in the means. A trend with two color changes represents differences among all three means. A dual colored charting point (red and green) represents a statistically significant difference between one mean but not the other. The reader is encouraged to examine the individual table of means for these cases.

In terms of grade level equivalents, students are expected to progress at least one grade level per year. Over the two complete calendar years, the average progress of students in Cohort A was 1.1 grade equivalents in language, 2 grade equivalents in math/problem solving, and 1.6 grade equivalents in reading. Thus, it seems that these students were making the expected progress only in math, but falling further behind the national comparison sample of the MAT-7 in language arts and reading.

In Cohort B, students progressed 1.8 grade equivalents in language, 2.1 grade equivalents in math/problem solving, and 1.6 grade equivalents in reading. The pattern in both cohorts was similar, indicating that the students were not progressing as quickly as might be expected in language and reading, although they were progressing as expected in math over two academic years. On the standard score scale, students in both cohorts increased each year in each of the three subjects except Cohort A on the language subtest where there was only an overall statistically significant gain from 1997 to 1999.

Although the grade equivalents and standard score analyses indicate that the students are progressing each year, the national percentile and the normal curve equivalent score analyses allow us to compare the growth of this cohort of students with national norms. The students in Cohort A failed to make significant gains relative to these national norms. In language, there was actually a decreasing trend in the mean national percentile scores, although not statistically significant. Mean NCEs, however, did show a statistically significant decrease in language for Cohort A between 1997 and 1999.

Cohort B performed slightly better relative to the national norms for mean national percentile and normal curve equivalents. There were no statistically significant gains for either language or reading, but there was a statistically significant gain on the math subtest between 1997 and 1998 (6<sup>th</sup> grade to 7<sup>th</sup> grade).

## SAT-9

Since only two years of data were made available to us on the SAT-9, our analyses and conclusions are limited. SAT-9 data was provided on three subtests in four scales: language, mathematics, and reading in grade equivalent (GE), standard scores (SS), percentile rank (PR), and National Curve Equivalent (NCE). We defined two cohorts based on the available data; Cohort C constituted those students at Seven Hills who progressed from third to fourth grade from 1997/98 to 1998/99, and Cohort D included those students in Seven Hills who progressed from fourth to fifth grade from 1997/98 to 1998/99. Sample sizes varied slightly over the subtests. In Cohort C, there were 63, 62, and 58 students in the language, mathematics and reading subtest analyses respectively. In Cohort D there were 71, 72, and 72 students in the language, mathematics, and reading subtest analyses respectively.

## Longitudinal analysis findings

Dependent t-test results for both cohorts indicated a clear pattern of increasing performance for the Seven Hills students. All pairwise differences were statistically significant except for Cohort C in reading on the PR and NCE scales. Table 8:2 presents a summary of these findings.

Table 8:2 Dependent t-Test Summary Findings for Cohort C and D

	Cohort C		Cohort D	
	Mean Diff	Std	Mean Diff	Std
GE Lang	2.56*	2.21	1.58*	2.91
GE Math	1.97*	1.87	2.45*	1.98
GE Reading	1.36*	1.95	1.23*	1.48
SS Lang	38.52*	27.40	21.89*	33.99
SS Math	46.34*	31.25	43.35*	29.44
SS Reading	25.52*	28.80	23.43*	23.94
PR Lang	12.76*	17.69	8.68*	22.26
PR Math	16.40*	21.04	16.89*	22.00
PR Reading	2.02	18.56	5.19*	17.79
NCE Lang	9.46*	13.94	5.46*	16.37
NCE Math	11.01*	15.42	11.09*	15.25
NCE Math	1.59	14.20	3.91*	11.94

\*  $p < .05$

## 8.4 Chi-Square Analysis of MCAS Data

A chi-square analysis was initiated on data available from the state of Massachusetts on the outcomes of the Massachusetts Comprehensive Assessment Systems (MCAS), the state-mandated criterion-referenced test. While the chi-square can help us distinguish the size and strength of the differences between Seven Hills Charter School and the two control groups we compared it with, our main interest is in the change over time at this Edison school. The chi-square analysis was completed because it is one step in preparation for the odds ratio analysis, which allows us to examine relative changes over time in this school or in the control groups.

The MCAS is administered in grade 4 (English, mathematics, and science and technology) and grade 8 (English, mathematics, history, and science and technology). We did not include grade 8 history, since there was only one year of data available. The MCAS is scored along an ordinal four category scale: Fail, Needs Improvement, Proficient, and Advanced. Additional information on the MCAS is provided in Appendix A.

## Construction of the comparison groups

We constructed two different comparison groups for the chi-square analyses. Since we were interested in examining the number/proportion of students who met state standards (“passing”) or conversely the number/proportion of students who did not meet state standards (“failing”) on the MCAS, we needed to define a suitable comparison group. Our first comparison was with the local public school district in which Seven Hills Charter School resides (i.e., Worcester Public Schools). The second comparison group we selected was the state average passing/failing rates. While the state demographics differ from Seven Hills, we believe that comparisons with state averages can yield information regarding the relative gains of this Edison school. Also, since Edison claims that advances in other district schools are—in part—due to its presence, we use the state as a more distant point of comparison that cannot be easily influenced by the presence of Edison schools.

## General procedure

Utilizing published data from the state of Massachusetts, we made comparisons over two years (Spring 1998 and Spring 1999) for both grade 4 and grade 8. Percentage data (students in each scoring category) were converted to raw frequency data prior to chi-square analysis. To insure independence of the rows in the chi-square tables, the raw frequencies for each scoring category of the MCAS in the district and state comparisons were down-weighted by subtracting the number of students in that category from Seven Hills. Thus, both the district and state numbers reflect all students in the district or state exclusive of those in Seven Hills Charter School.

Four chi-square analyses were evaluated for each subtest nested within year and grade level. Two of these analyses were on uncollapsed data; that is, all scoring levels were represented in the contingency table (e.g., a 2x4) for the district and state comparisons. Two follow-up analyses were conducted on the data after collapsing the multilevel scoring into a dichotomy (pass, fail), thus producing 2x2 contingency tables. According to the Massachusetts Department of Education, a score in the “proficient” or “advanced” category constitutes “passing,” or meeting the state standards for that particular grade and subject. On the other hand, the “fail” and “needs improvement” categories mean that the student has not met the state passing standards. Thus, for the collapsed chi-square and odds ratio analyses, we grouped the “proficient” and “advanced” levels into a passing or meeting state standards and the “needs improvement” and “failing” categories into failing or not meeting state standards.

## Chi-square findings

The chi-square analyses test the null hypothesis that the relative frequency (of students) in the four (or two) scoring categories are the same for Seven Hills and the comparison group (either the district or the state). Results of the chi-square analyses for grade 4 covered the MCAS administration for the

two years from 1998 to 1999 for English, mathematics, and science and technology. Individual contingency tables are presented in Appendix E.

The first set of comparisons was made against district data as the comparison group. Six separate chi-square statistics were evaluated from 2x4 contingency tables and six from 2x2 contingency tables (see Table 8:3) for each year. Results of the three subtests administered in grade 4 closely parallel each other. Statistically significant differences were observed relative to the state proportions and district proportions in each year and subtest. Fewer students fell into the “proficient” and “advanced” categories in Seven Hills relative to the overall state and district. This pattern was also replicated in the collapsed analyses (2x2) except on the 1998 English and 1999 mathematics tests.

Results of the grade 8 MCAS chi-square analysis are summarized in Table 8:4. Similar to the grade 4 chi-square analyses, the analyses for grade 8 covered the MCAS administration for the two years from 1998 to 1999 for English, mathematics, and science and technology. Individual contingency table results are presented in Appendix E. The first set of comparisons was made against district data as the comparison group. Six separate chi-square statistics were evaluated from 2x4 contingency tables and six from 2x2 contingency tables (see Table 8:4) for each year. Results among the three subtests taken at grade 8 closely paralleled one another in 1998. In general, students at Seven Hills performed at levels commensurate with students at the district

Table 8:3 Summary of Chi-Square Findings for Seven Hills Charter School, Grade 4

	1998	1999
<i>English</i>		
Seven Hills vs. District	sig/ns	sig/sig
Seven Hills vs. State	sig/sig	sig/sig
<i>Mathematics</i>		
Seven Hills vs. District	sig/sig	sig/ns
Seven Hills vs. State	sig/sig	sig/sig
<i>Science &amp; Technology</i>		
Seven Hills vs. District	sig/sig	sig/sig
Seven Hills vs. State	sig/sig	sig/sig

Note: Each result cell in the matrix is divided with the results for the 2x4 analysis on the left-hand side and results for 2x2 analysis on the right-hand side.

Table 8:4 Summary of Chi-Square Findings for Seven Hills Charter School, Grade 8

	1998	1999
<i>English</i>		
Seven Hills vs. District	ns/ns	ns/sig
Seven Hills vs. State	ns/sig	sig/sig
<i>Mathematics</i>		
Seven Hills vs. District	ns/ns	sig/sig
Seven Hills vs. State	sig/ns	sig/sig
<i>Science &amp; Technology</i>		
Seven Hills vs. District	ns/ns	ns/sig
Seven Hills vs. State	ns/sig	sig/sig

Note: Each result cell in the matrix is divided with the results for the 2x4 analysis on the left-hand side and results for 2x2 analysis on the right-hand side.

and state except in three analyses. In these three exceptions, a greater proportion of students were failing at Seven Hills than in the comparison group.

## 8.5 Odds Ratio Analysis of the MCAS Data

One of the many possible statistics that can be derived from a 2x2 contingency table is the odds ratio statistic (OR) and corresponding  $1-\alpha$  confidence interval. As presented in Section 2.4 of this report, the 2x2 tables analyzed in the previous section can be thought of as representing consecutive class cohorts in a prospective design. From a classical epidemiological perspective, the students in the Edison school can be thought of as the “exposed” group, that is, exposed to the “Edison-effect,” and students in the comparison group as the unexposed group. From this perspective, each yearly comparison is a new cohort, measured over a period of years. There is a minimal possibility for cohort contamination if a number of students in one group are not promoted to the next grade level. However, we think this represents a very small number of possible cases and therefore has minimal impact on the validity of these analyses. Section 2.4 details the OR statistic and corresponding  $1-\alpha$  confidence interval.

We calculated and charted OR for each of the 2x2 tables constructed from the chi-square analyses presented above. Tables 8:5 and 8:6 and Exhibits 8:3 and 8:4 present these findings. Exhibits 8:3 and 8:4 graphically illustrate the overall performance of the Seven Hills Charter School on the MCAS in comparison with district and state performance levels at the 4<sup>th</sup> and 8<sup>th</sup> grades, respectively. The charts in Exhibits 8:5 - 8:11 illustrate the distribution of scores across all four performance categories. It is important to remember that the odds ratio findings consider only two categories, with “Advanced” and “Proficient” combined for the pass category and “Needs Improvement” and “Fail” combined for the fail category in the odds ratio analysis.

### Odd ratio findings, grade 4

**Comparison against the district.** Three OR analyses were evaluated, one for each subject test on the MCAS. In grade 4 English, the OR for a Seven Hills student failing the 1998 English MCAS relative to the students in the district as a whole was 2.535; in 1999 it was 3.893. In other words, students attending Seven Hills Charter School were 2.5 times more likely to not meet state standards than students enrolled in other district schools in 1998. Since the odds ratio increased to 3.8 in 1999, this indicates that the performance at the Edison school on the MCAS was not as good as the performance of other grade 4 students in the district. The Breslow-Day chi-square for testing the hypothesis of homogeneity of OR over the two years was not statistically significant, indicating there was no real (statistically significant) change in OR. Thus, a common OR = 3.194 with a lower bound (LB) = 1.623 and an upper bound (UB) = 6.286 indicated that the Seven Hills’ grade 4 students were about three times more likely to fail (i.e., score in either the “needs improvement” or “fail” categories) the English MCAS than students in the rest of the district. The OR analysis of the grade 4 mathematics component of the MCAS presented a different picture. The Breslow-Day chi-square for testing the hypothesis of homogeneity of OR over the two years just barely missed conventional levels of statistically significant ( $p = .053$ ). If one were to interpret this that a common

OR cannot be reasonably interpreted, ORs need to be examined for each year. In 1998 the OR was statistically significant—OR = 3.421, LB = 1.709, UB = 6.846—indicating that Seven Hills’ students were much more likely to fail relative to students in the rest of the district. In 1999, however, the OR decreased to a nonstatistically significant value of 1.422. Thus, the students at Seven Hills showed a rather large improvement in passing the MCAS math test in 1999.

Regarding the MCAS subtest of science and technology, the Breslow-Day chi-square for testing the hypothesis of homogeneity of OR

over the two years was not statistically significant, indicating there was no real (statistically significant) change in OR. Thus, a common OR = 3.154 with LB = 2.175 and UB = 4.574 indicated that the Seven Hills students were more than three times as likely to fail this test relative to the rest of the district. Table 8:5 contains a summary of the odds ratio findings for grade 4.

**Comparison against state.** In grade 4 English, the OR for a Seven Hills student failing the 1998 English MCAS relative to students in the state as a whole was 4.882, and in 1999 it was 5.256. In both years the 1- $\alpha$  CI did not eclipse 1 and thus can be interpreted as statistically significant. The Breslow-Day chi-square for testing the hypothesis of homogeneity of OR over the two years was not statistically significant, indicating there was no real (statistically significant) change in OR. Thus, a common OR = 5.067 with LB = 2.677 and UB = 9.591 indicated that the Seven Hills’ students were about 5 times more likely to fail the MCAS English subtest as compared with students in the state as a whole.

The OR analysis of the grade 4 mathematics component of the MCAS presented a similar picture. The Breslow-Day chi-square for testing the hypothesis of homogeneity of OR over the two years was not statistically significant ( $p = .062$ ). Thus, the common OR was 2.919, LB = 1.954, UB = 4.361. Although students at Seven Hills showed a rather large improvement on OR, they were still about three times more likely to fail the math MCAS.

Regarding the MCAS subtest of science and technology, the Breslow-Day chi-square for testing the hypothesis of homogeneity of OR over the two years was not statistically significant, indicating there was no real (statistically significant) change in OR. Thus, the common OR = 4.525 with LB = 3.191 and UB = 6.416 indicated that the Seven Hills students had a substantially greater odds for failing (about 4½ times) than students in the rest of the state.

Table 8:5 Summary of Odds Ratio Findings for Seven Hills Charter School, Grade 4

	1998	1999	2-year OR
<i>Odds of not meeting standard compared with district</i>			
English	2.535	3.893	3.194
Mathematics	3.421	1.422	2.045
Science & Tech.	2.799	3.572	3.154
<i>Odds of not meeting standard compared with state</i>			
English	4.882	5.256	5.067
Mathematics	4.763	2.065	2.919
Science & Tech.	3.588	5.779	4.525

## Odds ratio findings, grade 8

**Comparison against the district.** In grade 8, the OR for Seven Hills' students showed an increasing, but not statistically significant, trend for odds of failure on the English component of the MCAS relative to students in the district. The Breslow-Day chi-square for testing the hypothesis of homogeneity of OR over the two years was not statistically significant, indicating that a common OR over the two years was meaningful, OR = 1.417, LB = 0.967, UB = 2.076. Thus, for the two-year period the Seven Hills' students were at even odds for failure relative to students in the rest of the district.

On the math subtest, students at Seven Hills showed a statistically significant rise in the OR, e.g., the Breslow-Day chi-square for testing the hypothesis of homogeneity of OR over the two years was statistically significant. In 1998 the OR was 0.770 and rose to 3.485 in 1999.

The grade 8 MCAS subtest of science and technology evidenced a pattern similar to math. The Seven Hills' students started out in 1998 at even odds relative to district students but fell behind in 1999, OR = 2.729, LB = 1.132, UB = 6.580. Table 8:6 contains a summary of the odds ratio findings for grade 8.

**Comparison against state.** In grade 8, the ORs for Seven Hills' students revealed an increasing odds for failure on both the English and math MCAS relative to other students in the state. In both analyses the Breslow-Day chi-square for testing the hypothesis of homogeneity of OR over the two years was statistically significant. For English, the 1998 OR = 1.863, LB = 1.084, UB = 3.204, indicated that Seven Hills' students were slightly more likely to fail. However, in 1999 the OR = 4.566, LB = 2.834, UB = 7.357, indicating that Seven Hills' students lost ground relative to students in the rest of the state. On the math subtest, students at Seven Hills showed a statistically significant rise in the OR, e.g., in 1998 the OR=1.741 and rose to 7.503 in 1999.

Regarding the grade 8 MCAS subtest of science and technology, the Breslow-Day chi-square for testing the hypothesis of homogeneity of OR over the two years was not statistically significant, indicating there was no real (statistically significant) change in OR. Thus, a common OR = 3.593 with LB = 2.104 and UB = 6.135 indicated that the Seven Hills' students were about 3 ½ times more likely to fail the science and technology MCAS relative to students in the rest of the state.

Table 8:6 Summary of Odds Ratio Findings for Seven Hills Charter School, Grade 8

	1998	1999	2-year OR
<i>Odds of not meeting standard compared to district</i>			
English	1.034	1.828	1.417
Mathematics	0.770	3.485	* p=.010
Science & Tech.	0.085	2.729	* p=.046
<i>Odds of not meeting standard compared to state</i>			
English	1.863	4.566	* p=.019
Mathematics	1.741	7.503	* p=.011
Science & Tech.	2.189	5.840	3.593

Note: The asterisks indicate that the difference in OR over two years was statistically significant, thus making it impossible to calculate an overall OR. The p-value is listed in place of the 2-year OR.

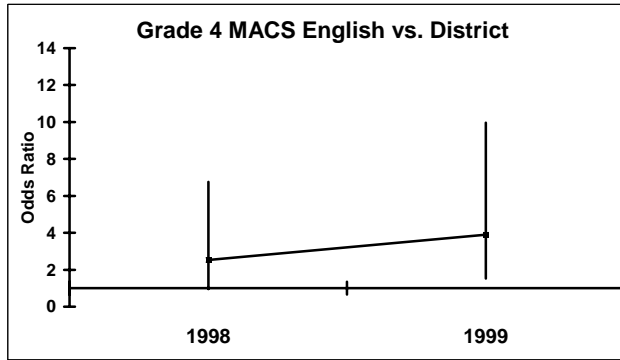
### Exhibit 8:3 Results of the Odds Ratio Analysis for Seven Hills Charter School (Grade 4)

#### Grade 4 MCAS English

Year	U CI	L CI	OR
1998	6.756	0.951	2.535
1999	9.966	1.521	3.893

Breslow-Day for Homogeneity of Odd Ratio  
Chi-Sq (1, N=3913) = 0.347, p = .556

OR = 3.194  
UB = 6.286  
LB = 1.623

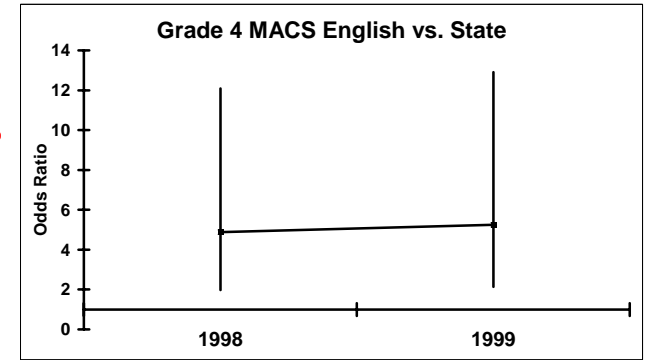


#### Grade 4 MCAS English

Year	U CI	L CI	OR
1998	12.088	1.971	4.882
1999	12.902	2.141	5.256

Breslow-Day for Homogeneity of Odd Ratio  
Chi-Sq (1, N=151,335) = 0.10, p = .919

OR = 5.067  
UB = 9.591  
LB = 2.677

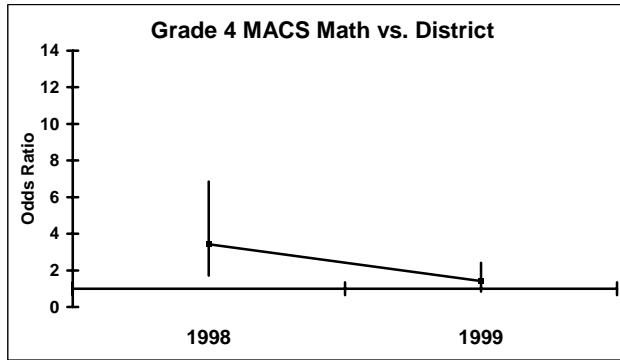


#### Grade 4 MCAS Math

Year	U CI	L CI	OR
1998	6.846	1.709	3.421
1999	2.413	0.838	1.422

Breslow-Day for Homogeneity of Odd Ratio  
Chi-Sq (1, N=4,032) = 3.736, p = .053

OR = 2.045  
UB = 3.105  
LB = 1.348

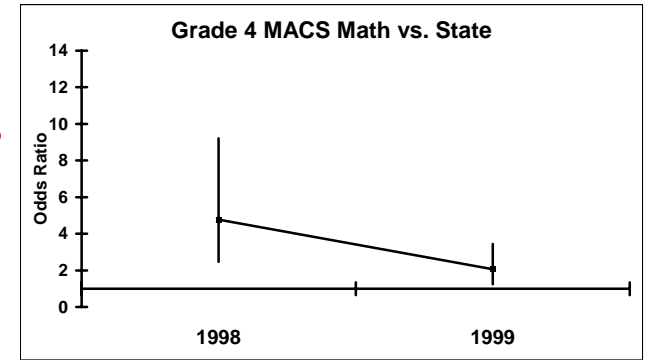


#### Grade 4 MCAS Math

Year	U CI	L CI	OR
1998	9.217	2.461	4.763
1999	3.442	1.239	2.065

Breslow-Day for Homogeneity of Odd Ratio  
Chi-Sq (1, N=152,225) = 3.472, p = .062

OR = 2.919  
UB = 4.361  
LB = 1.954

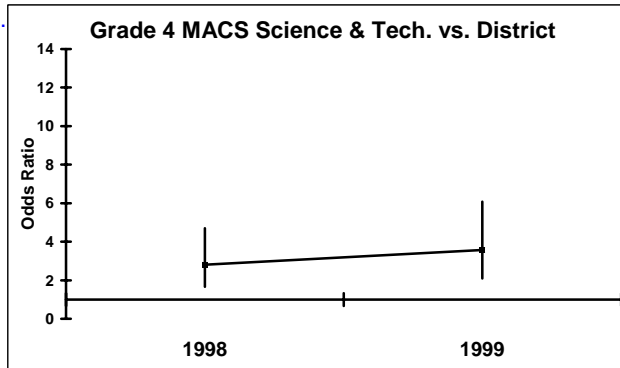


#### Grade 4 MCAS Science & Tech.

Year	U CI	L CI	OR
1998	4.707	1.665	2.799
1999	6.085	2.097	3.572

Breslow-Day for Homogeneity of Odd Ratio  
Chi-Sq (1, N=4,032) = 0.373, p = .542

OR = 3.154  
UB = 2.175  
LB = 4.574

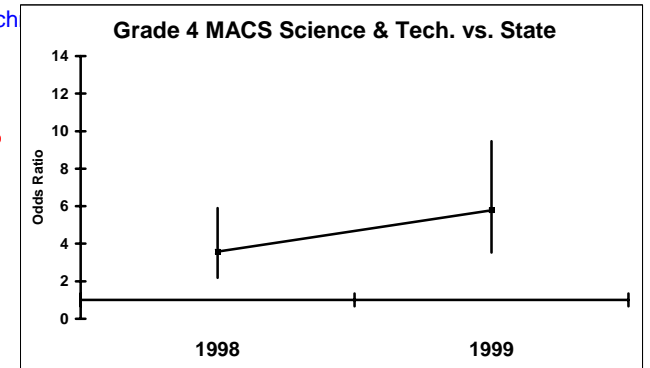


#### Grade 4 MCAS Science & Tech.

Year	U CI	L CI	OR
1998	5.908	2.179	3.588
1999	9.469	3.528	5.779

Breslow-Day for Homogeneity of Odd Ratio  
Chi-Sq (1, N=152,234) = 1.469, p = .225

OR = 4.525  
UB = 6.416  
LB = 3.191



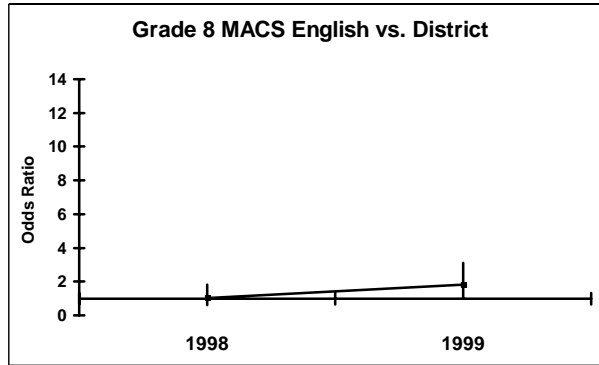
### Exhibit 8:4 Results of the Odds Ratio Analysis for Seven Hills Charter School (Grade 8)

#### Grade 8 MCAS English

Year	U CI	L CI	OR
1998	1.811	0.591	1.034
1999	3.097	1.079	1.828

Breslow-Day for Homogeneity of Odd Ratio  
Chi-Sq (1, N=3,189) = 2.100, p = .147

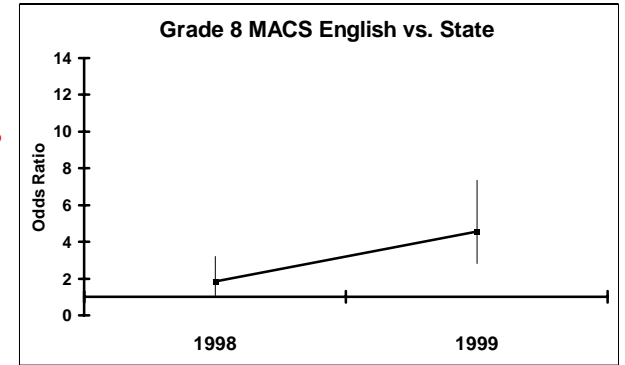
OR = 1.417  
UB = 2.076  
LB = 0.967



#### Grade 8 MCAS English

Year	U CI	L CI	OR
1998	3.204	1.084	1.863
1999	7.357	2.834	4.566

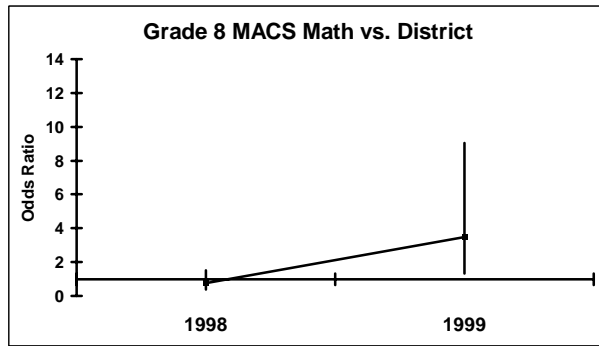
Breslow-Day for Homogeneity of Odd Ratio  
Chi-Sq (1, N=138,021) = 5.468, p = .019



#### Grade 8 MCAS Math

Year	U CI	L CI	OR
1998	1.520	0.394	0.770
1999	9.046	1.343	3.485

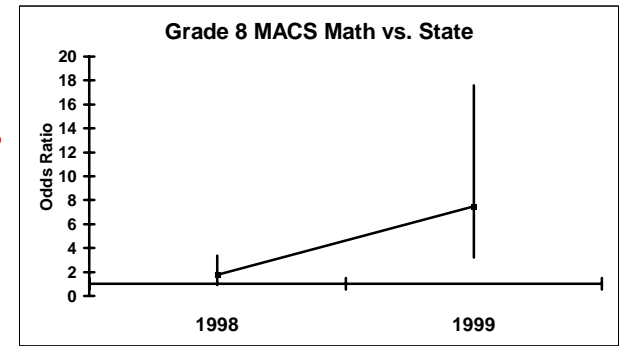
Breslow-Day for Homogeneity of Odd Ratio  
Chi-Sq (1, N=3,275) = 6.604, p = .010



#### Grade 8 MCAS Math

Year	U CI	L CI	OR
1998	3.354	0.904	1.741
1999	17.620	3.195	7.503

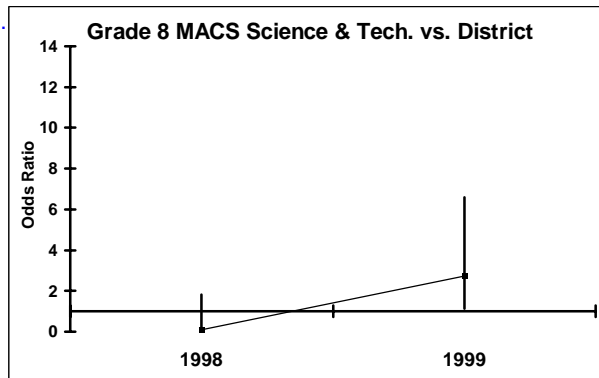
Breslow-Day for Homogeneity of Odd Ratio  
Chi-Sq (1, N=138,937) = 6.434, p = .011



#### Grade 8 MCAS Science & Tech.

Year	U CI	L CI	OR
1998	1.818	0.393	0.085
1999	6.580	1.132	2.729

Breslow-Day for Homogeneity of Odd Ratio  
Chi-Sq (1, N=3,231) = 3.985, p = .046



#### Grade 8 MCAS Science & Tech.

Year	U CI	L CI	OR
1998	4.557	1.051	2.189
1999	12.970	2.630	5.840

Breslow-Day for Homogeneity of Odd Ratio  
Chi-Sq (1, N=140,306) = 2.830, p = .093

OR = 3.593  
UB = 6.135  
LB = 2.104

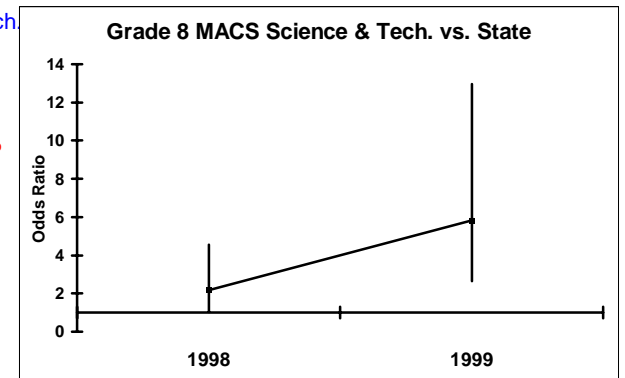
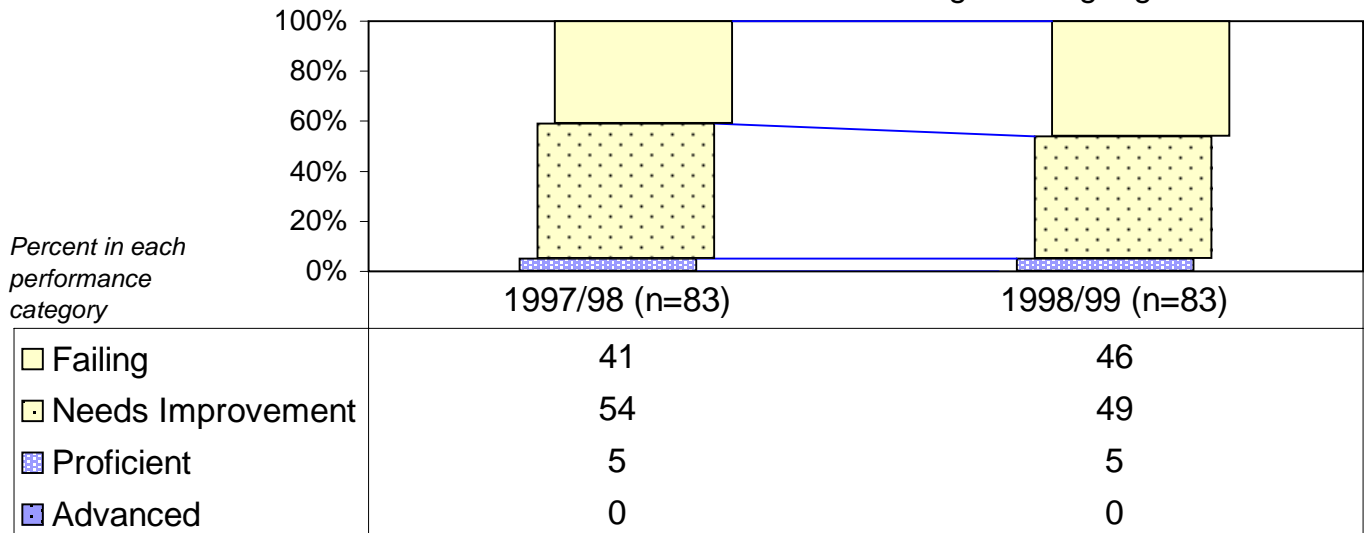
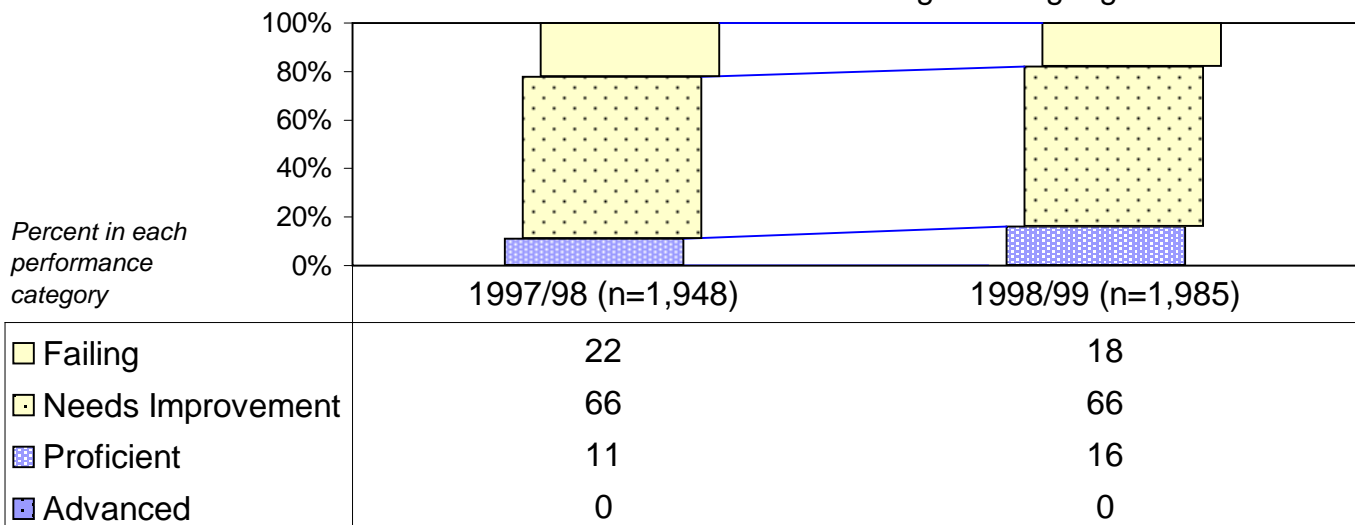


Exhibit 8:5 Performance on Grade 4 English Lang. Arts for Edison, District, and State

Seven Hills Charter School MCAS Results for 4th Grade English Language Arts



Worcester Public Schools MCAS Results for 4th Grade English Language Arts



State of Massachusetts MCAS Results for 4th Grade English Language Arts

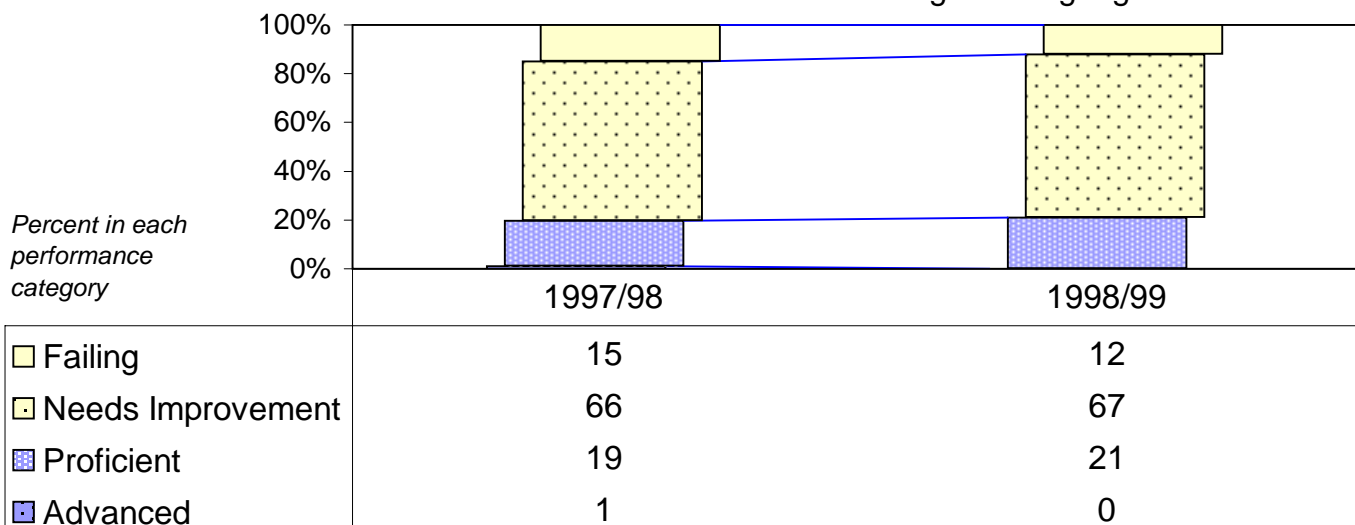
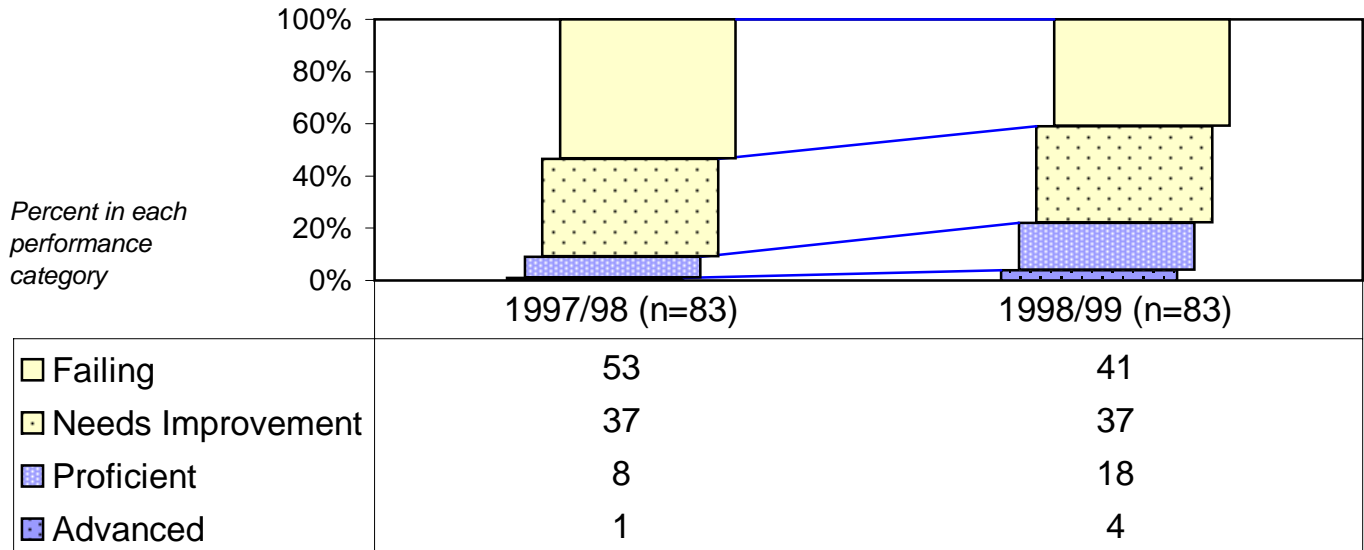
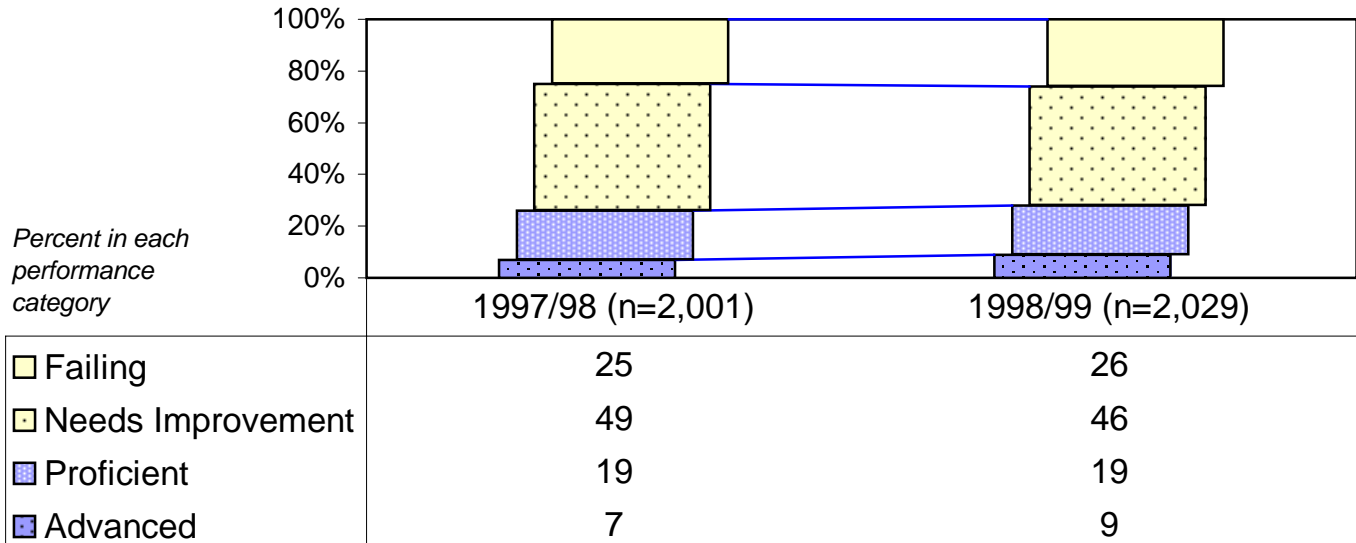


Exhibit 8:6 Performance on Grade 4 Math for Edison, District, and State

Seven Hills Charter School MCAS Results for 4th Grade Math



Worcester Public Schools MCAS Results for 4th Grade Math



State of Massachusetts MCAS Results for 4th Grade Math

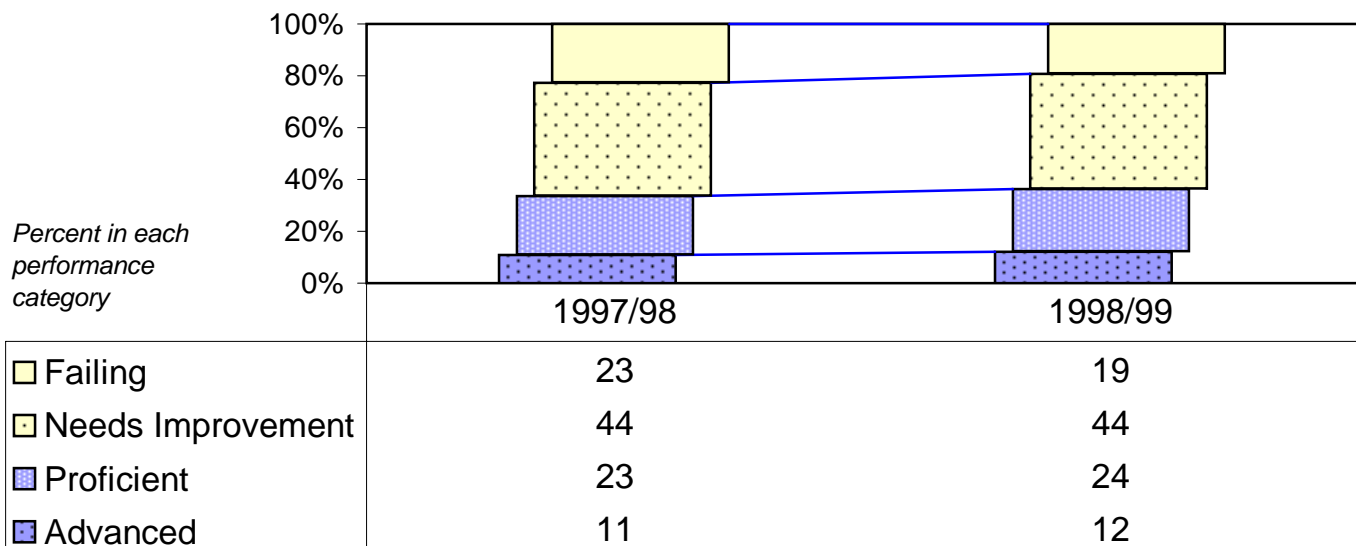
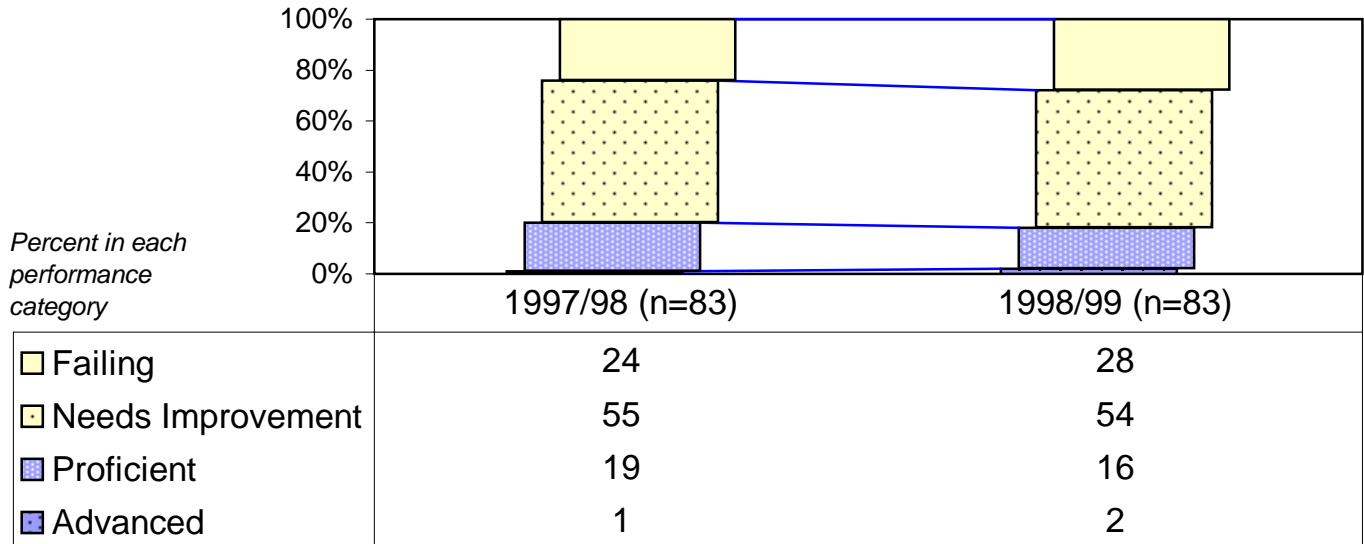
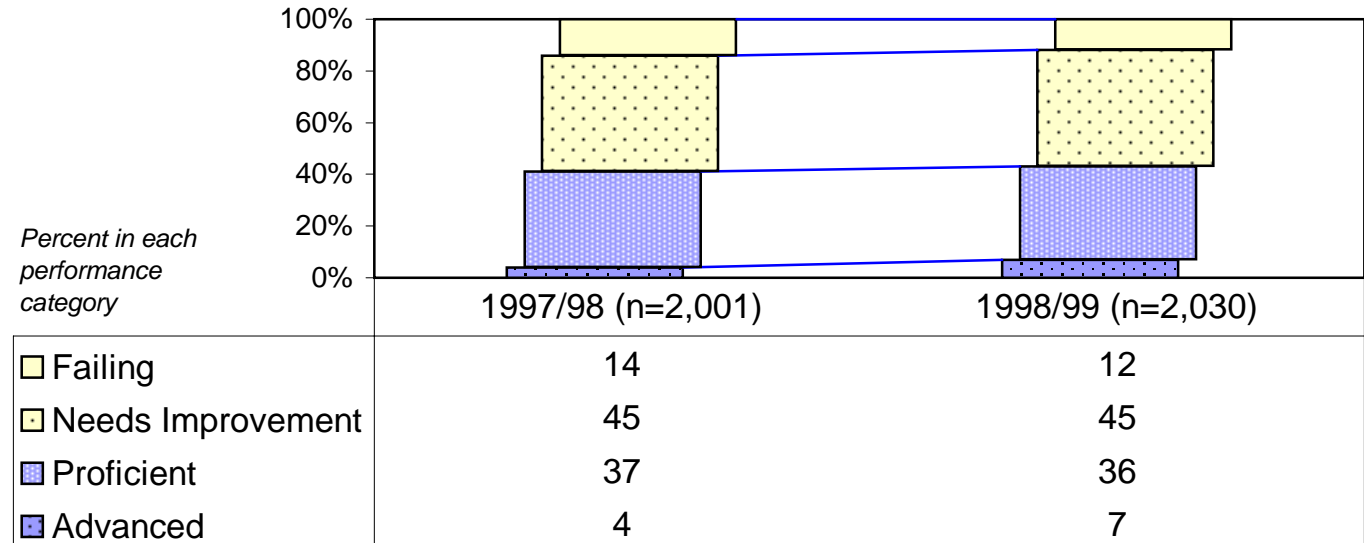


Exhibit 8:7 Performance on Grade 4 Science & Tech. for Edison, District, and State

Seven Hills Charter School MCAS Results for 4th Grade Science & Tech



Worcester Public Schools MCAS Results for 4th Grade Science & Tech.



State of Massachusetts MCAS Results for 4th Grade Science & Tech.

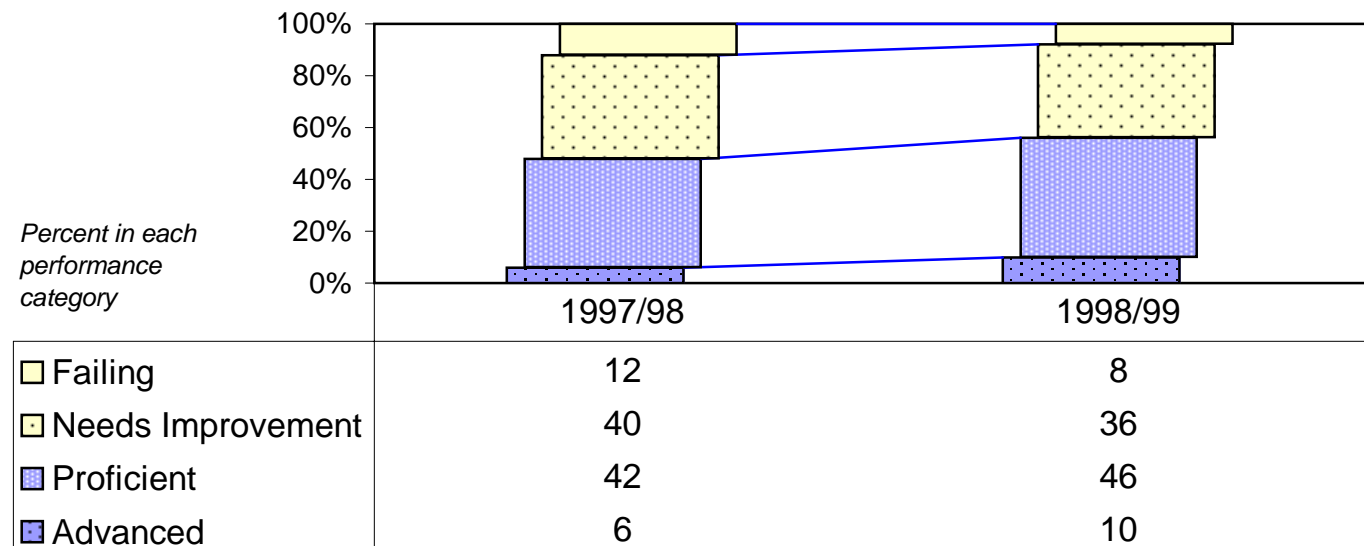
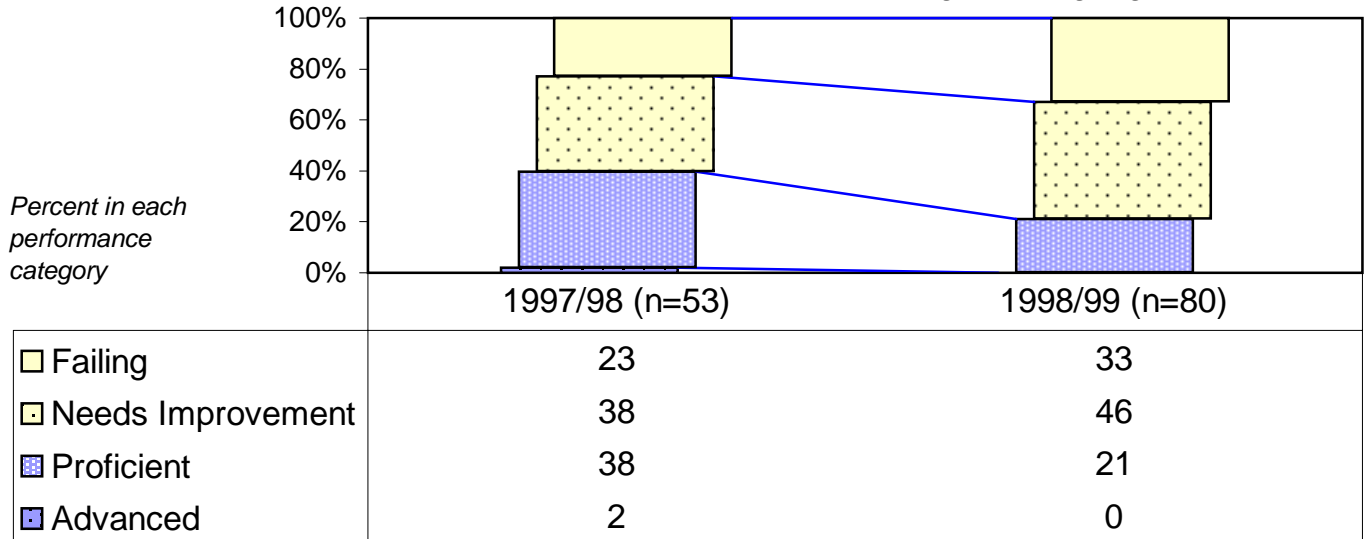
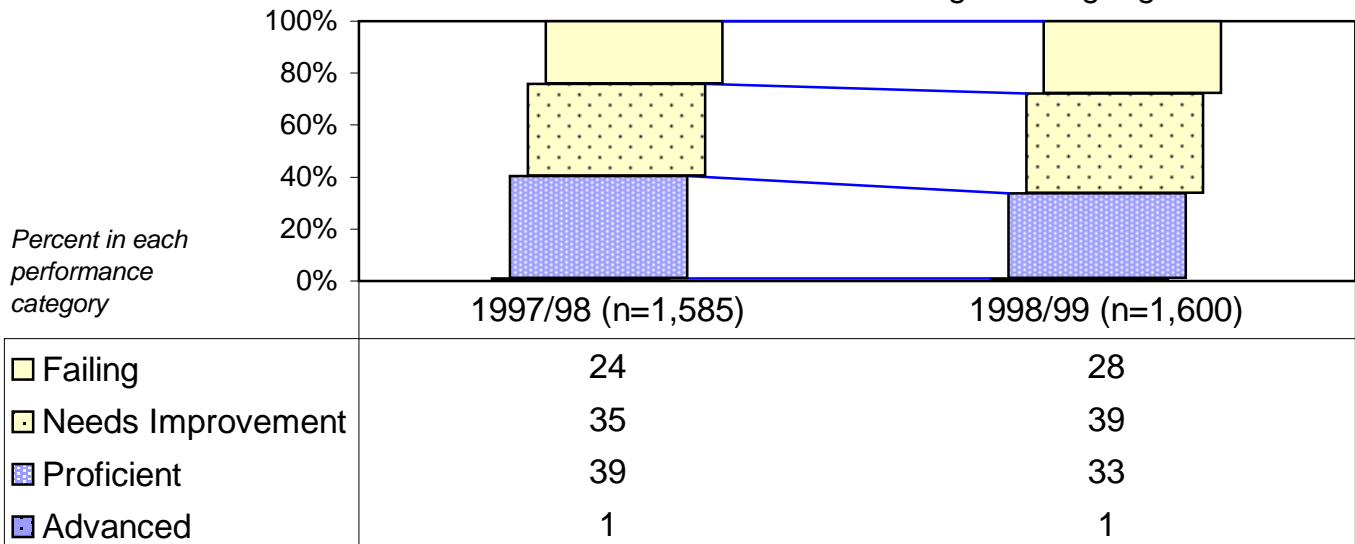


Exhibit 8:8 Performance on Grade 8 English Lang. Arts for Edison, District, and State

Seven Hills Charter School MCAS Results for 8th Grade English Language Arts



Worcester Public Schools MCAS Results for 8th Grade English Language Arts



State of Massachusetts MCAS Results for 8th Grade English Language Arts

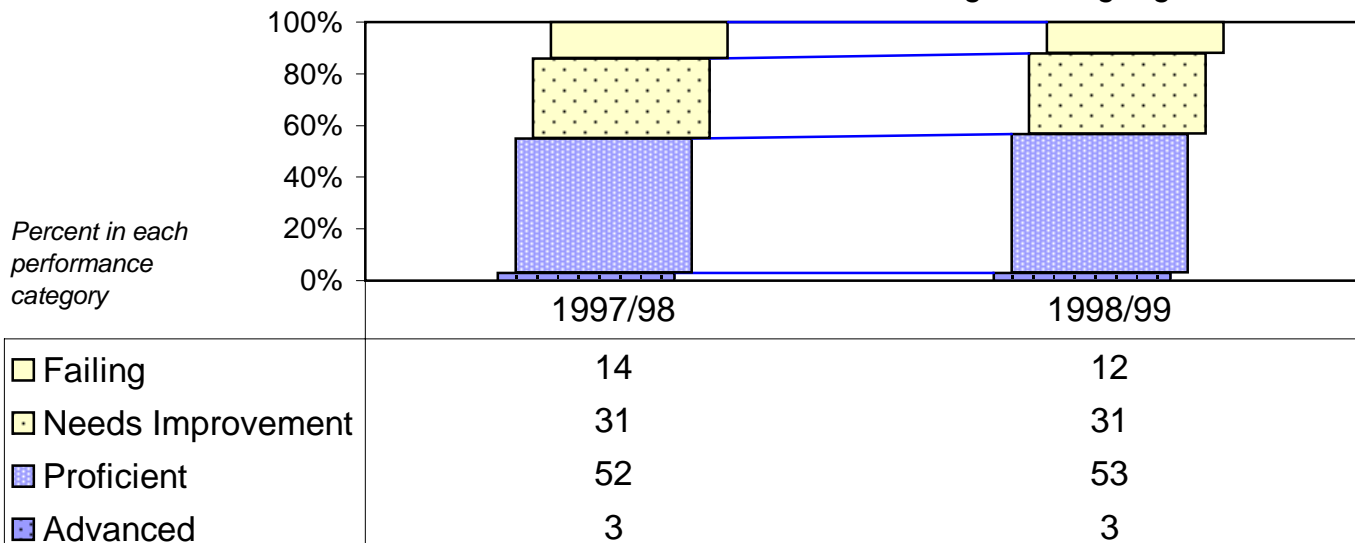
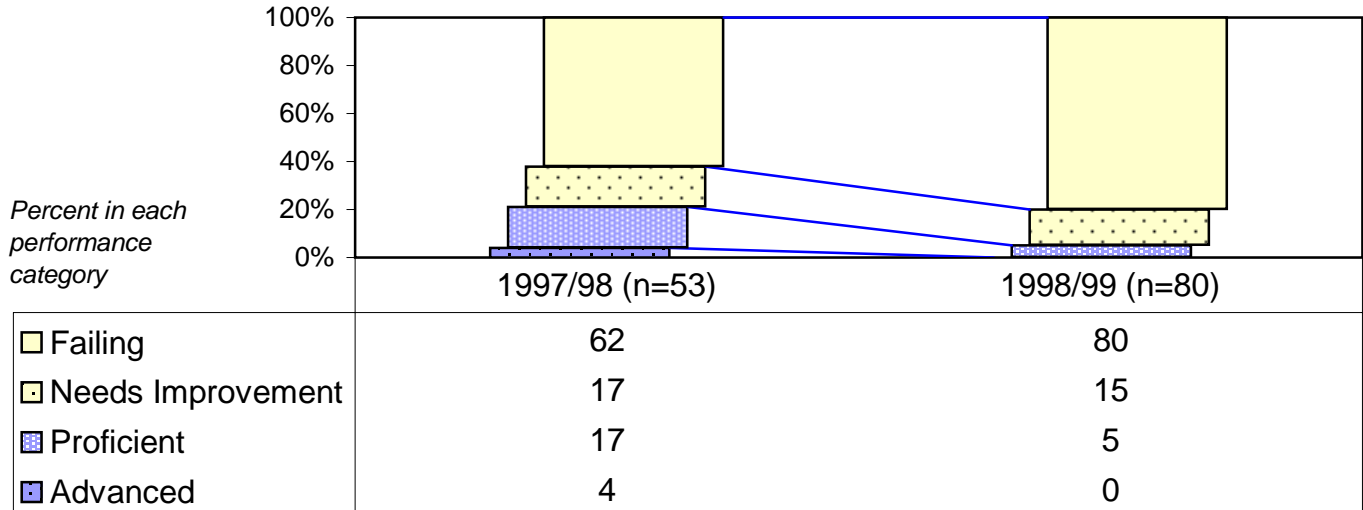
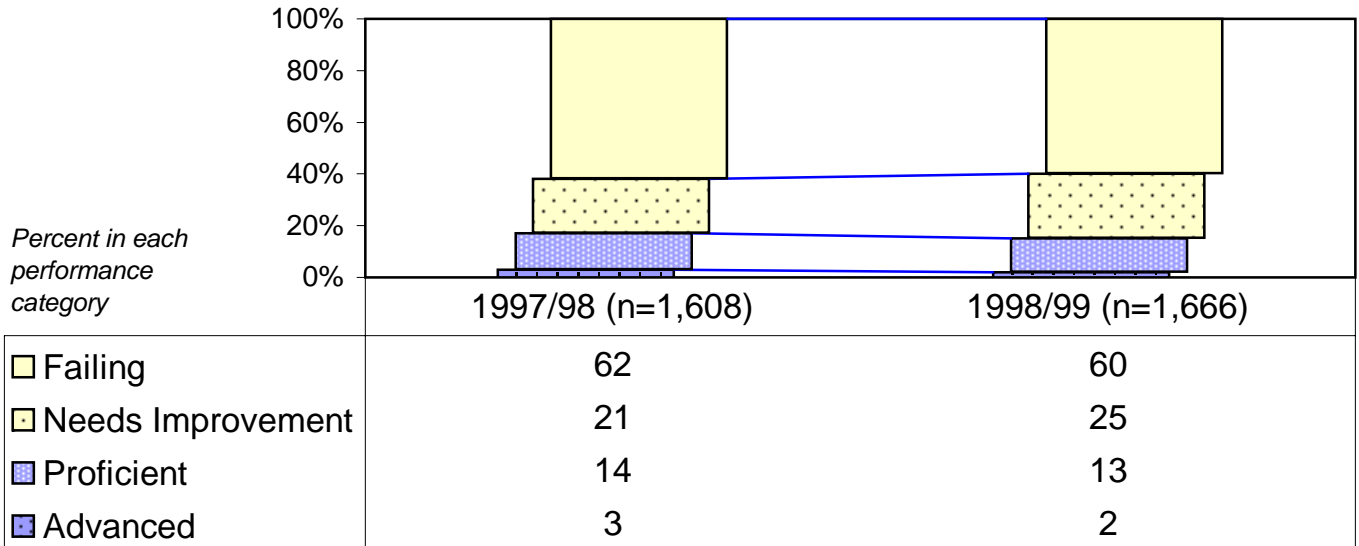


Exhibit 8:9 Performance on Grade 8 Math for Edison, District, and State

Seven Hills Charter School MCAS Results for 8th Grade Math



Worcester Public Schools MCAS Results for 8th Grade Math



State of Massachusetts MCAS Results for 8th Grade Math

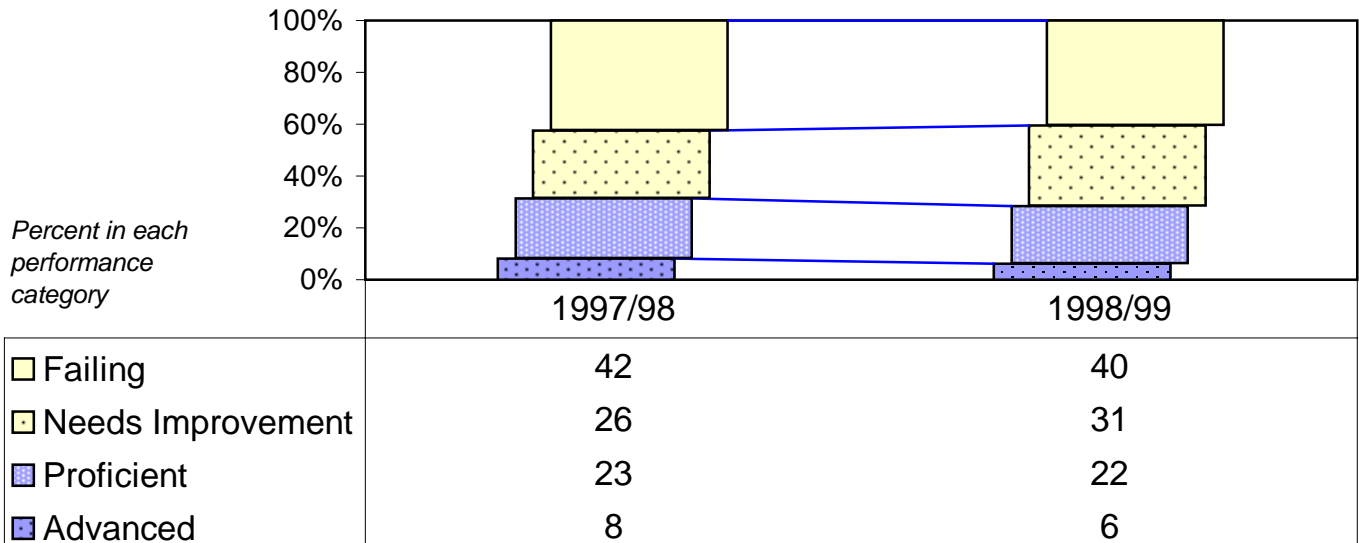
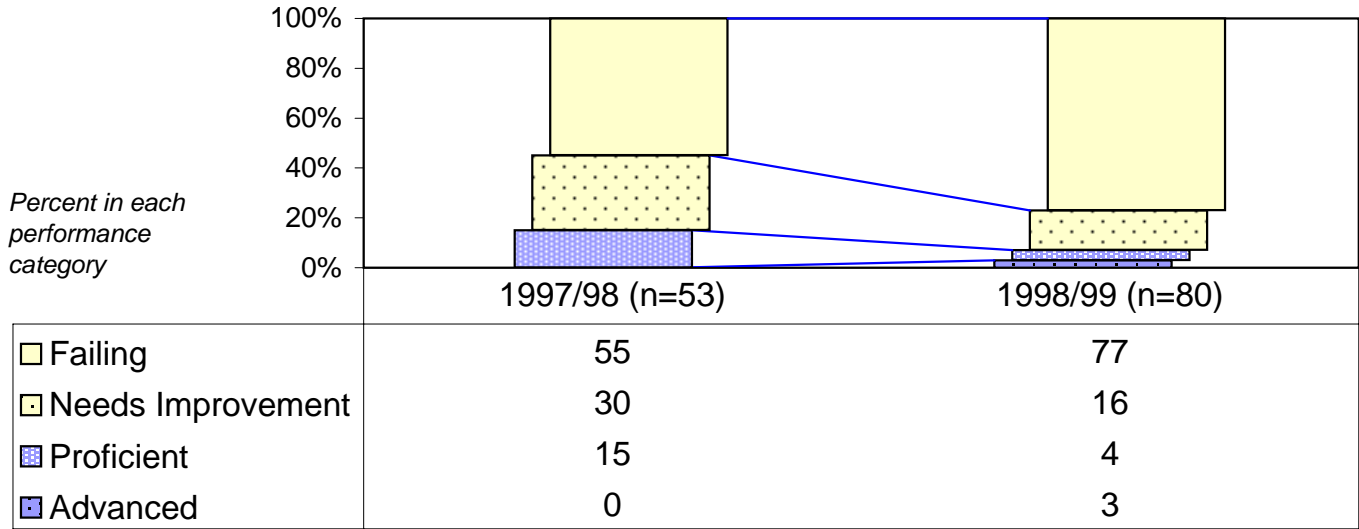
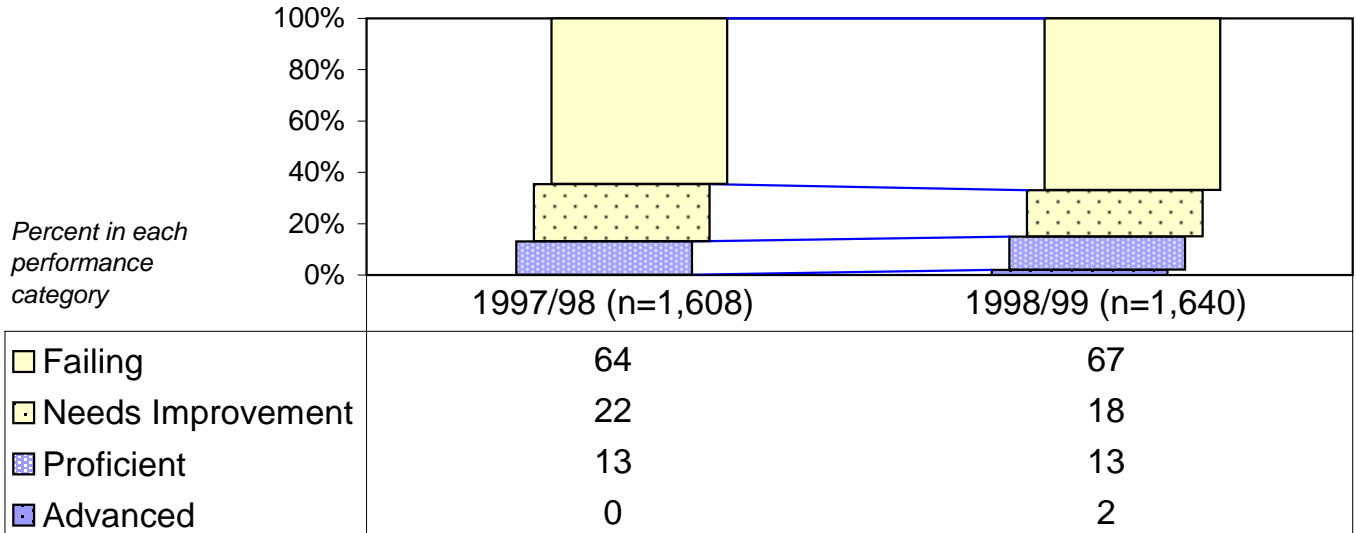


Exhibit 8:10 Performance on Grade 8 Science & Tech for Edison, District, and State

Seven Hills Charter School MCAS Results for 8th Grade Science & Tech



Worcester Public Schools MCAS Results for 8th Grade Science & Technology



State of Massachusetts MCAS Results for 8th Grade Science & Technology

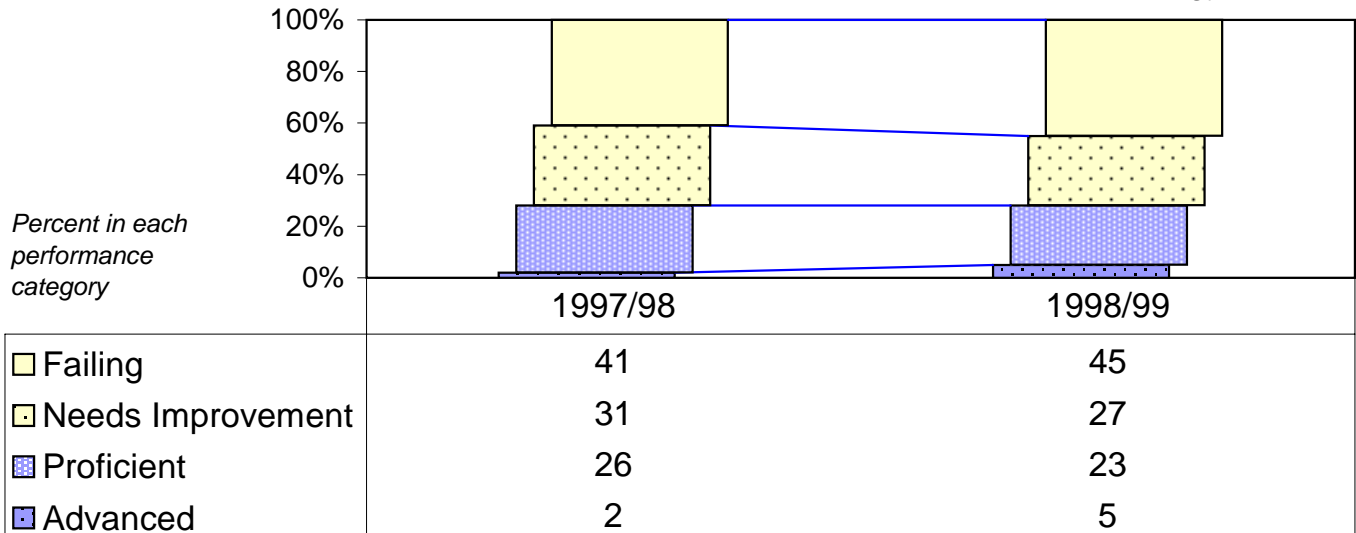
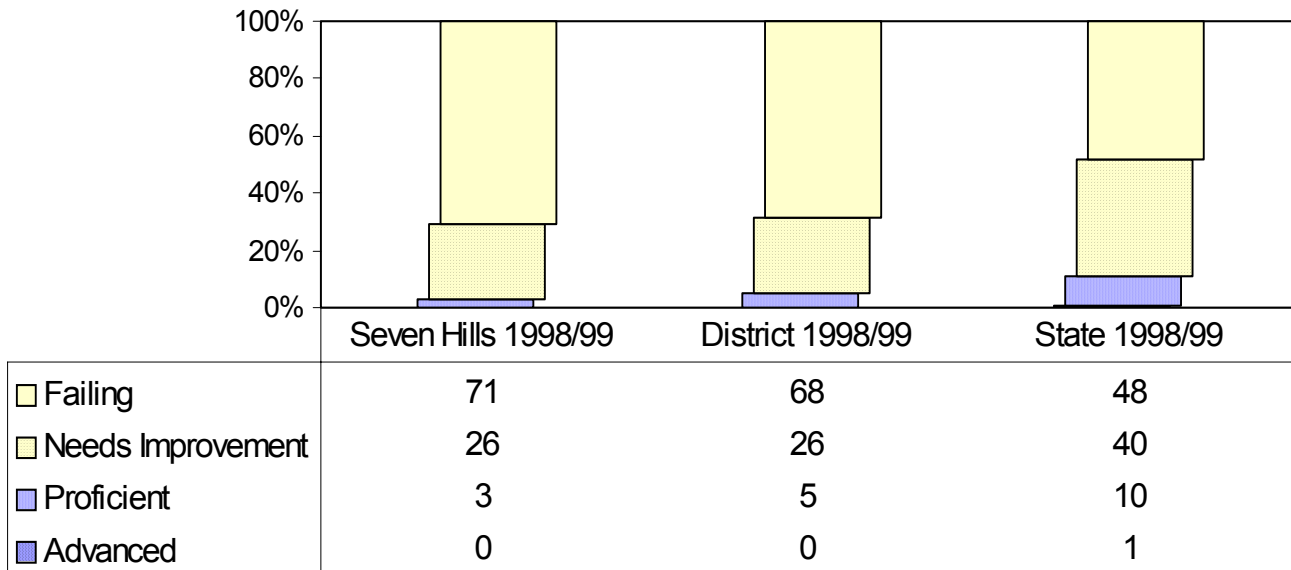


Exhibit 8:11 Performance on Grade 8 History and Social Studies for Edison, District, and State

MCAS Results for 8th Grade History and Social Studies



## 8.6 Summary

While the students, on the whole, made some gains on the norm-referenced test, the gain was mixed and often depended on the achievement test.

### Norm-referenced test findings

We constructed summary tables for aggregating our findings on the MAT-7 (Table 8:6) and SAT-9 (Table 8:7). A summary score of -1 indicates a result that is unfavorable toward the sample school, a score of 0 indicates a neutral finding, and a score of 1 indicates a favorable result according to the criteria specified in Section 2.5. The effect size (ES) is the omega squared ( $\omega^2$ ) for a one-way repeated measures ANOVA (Kepple, 1991).

Tables 8:6 and 8:7 highlight the main findings and trend ratings for the two norm-referenced test results considered. It appears that Seven Hills' students evidenced mild achievement growth over the three academic years, but did not make any noteworthy gains in comparison with the national norm. The Edison effect, as measured by the relative growth of student learning as measured by the MAT-7, indicated primarily no gains or losses relative to national norms over the three years we traced the students. Two exceptions to this are when the larger Cohort A showed a statistically significant decrease relative to the national norm for NCE over two years in language and the smaller Cohort B showed a statistically significant gain on the NCE over one year in math.

Table 8:6 Summary of Results on the Metropolitan Achievement Test (MAT-7)

<b>Cohort A</b> Grades 5,6,7 (1997-99)	Std Score		GE			PR			NCE			Trend
	p-value	ES	p-value	Δ	ES	p-value	Δ	ES	p-value	Δ	ES	
Language	.0005	.100	.0003	1.1	.110	.0788	-5.8	.024	.0476	<b>-4.7 (0)</b>	.0476	mixed (0)
Math	<.0001	.491	<.0001	2.0	.329	.4311	1.1	-.002	.3289	<b>1.3 (0)</b>	.002	mixed (0)
Reading	<.0001	.383	<.0001	1.6	.321	.2690	-2.6	.005	.5776	<b>-0.1 (0)</b>	-.007	mixed (0)
<b>Cohort B</b> Grades 6,7,8 (1997-99)	Std Score		GE			PR			NCE			Trend
	p-value	ES	p-value	Δ	ES	p-value	Δ	ES	p-value	Δ	ES	
Language	<.0001	.261	<.0001	1.8	.247	.4056	2.1	-.002	.3777	<b>2.4 (0)</b>	0.0	mixed (0)
Math	<.0001	.471	<.0001	2.1	.393	.0088	8.0	.108	.0167	<b>6.5 (0)</b>	.088	mixed (0)
Reading	<.0001	.330	<.0001	1.6	.359	.1515	2.4	.026	.3174	<b>1.9 (0)</b>	.005	mixed (0)

Table 8:7 Summary of Results on the Stanford Achievement Test (SAT-9)

<b>Cohort C</b> Grades 3&4 (1998-99)	Std Score		GE			PR			NCE			Trend
	p-value	ES	p-value	Δ	ES	p-value	Δ	ES	p-value	Δ	ES	
Language	<.0001	.495	<.0001	2.6	.400	<.0001	12.8	.201	<.0001	<b>9.5 (+1)</b>	.182	positive (+1)
Math	<.0001	.522	<.0001	1.9	.354	<.0001	16.4	.226	<.0001	<b>11.0 (+1)</b>	.198	positive (+1)
Reading	<.0001	.277	<.0001	1.4	.189	.4113	2.0	-.003	.3960	<b>1.6 (0)</b>	-.002	mixed (0)
<b>Cohort D</b> Grades 4&5 (1998-99)	Std Score		GE			PR			NCE			Trend
	p-value	ES	p-value	Δ	ES	p-value	Δ	ES	p-value	Δ	ES	
Language	<.0001	.167	<.0001	1.6	.123	.0016	8.7	.064	.0064	<b>5.5 (0)</b>	.046	mixed (0)
Math	<.0001	.519	<.0001	2.5	.430	<.0001	16.9	.223	<.0001	<b>11.1 (1)</b>	.205	positive (+1)
Reading	<.0001	.321	<.0001	1.2	.254	.0156	5.2	.034	.0070	<b>3.9 (0)</b>	.044	mixed (0)

### Criterion-referenced test findings

Decisions regarding the OR were based on whether or not the  $(1-\alpha)$  C.I. included 1.0. If the  $(1-\alpha)$  C.I. fell completely below 1.0, this was interpreted as a protective odds ratio (1), thus favoring the Edison School. If the  $(1-\alpha)$  C.I. included 1.0 (0), this was interpreted as an equal odds situation. If the  $(1-\alpha)$  C.I. fell completely above 1.0 (-1), this was interpreted as an increase in odds for failing the state CRT relative to the comparison sample.

Table 8:8 Summary Results on MCAS Criterion-Referenced Tests

Grade 4	1998	1999	B-D	Trend
District - English			-1	negative (-1)
District - Math			-1	negative (-1)
District - Science/Tech			-1	negative (-1)
Grade 8	1998	1999	B-D	Trend
District - English			0	mixed (0)
District - Math	0	-1		negative (-1)
District - Science/Tech	0	-1		negative (-1)

Note: All comparisons with the state as a comparison group were negative

If the Breslow-Day statistic (B-D) is nonsignificant, one overall OR and  $(1-\alpha)$  C.I can be used to represent the odds for failing the CRT relative to the comparison group. Thus, there are no trends reported for each specific year in the tables, only a rating in the B-D column. If the B-D statistic is found to be statistically significant, then an overall common OR cannot be meaningfully interpreted; that is, there is a statistically significant change in the OR over years and yearly ORs are necessary. Thus, our summary ratings appear for each year of data and not in the B-D column (see Table 8:8).

The odds ratio analysis of the MCAS results indicated that the odds of failing (i.e., scoring in the “needs improvement” or “failure” categories) the state assessment test were higher at Seven Hills Charter School than in the district and in the state as a whole. What is most disconcerting is that the odds of Seven Hills’ students failing increased over the two years in all subject areas but one in both grades 4 and 8. The one exception was grade 4 math where the odds of failing decreased at the Edison school compared with both district and state.

### Combined ratings

In terms of absolute scores, the school performed below national norms on the MAT-7 and below district and state levels on the state assessment test (MCAS). The overall achievement gains made by this school were slightly smaller than comparison groups.

Table 8:9 Combined Overall Trends for Seven Hills Charter School

	Positive	Mixed	Negative
Norm Referenced	3 of 12	9 of 12	0 of 12
Criterion Referenced	0 of 6	1 of 6	5 of 6
TOTALS	3 of 18	10 of 18	5 of 18

Table 8:9 contains the combined overall trends for Seven Hills Charter School. Our findings based upon the analysis of individual student results on the MAT-7, SAT-9, and the odds ratio analysis of consecutive cohorts on the MCAS, indicate that the performance of this school—in terms of student achievement—could best be characterized as Mixed with a mean rating of -0.11 because nearly all the trends were mixed although 3 NRT trends were positive and 5 of the CRT trends were moving in the wrong direction. In its 1999 annual report, Edison rated this school as Strongly Positive. In its 2000 annual report it rated the 1999-2000 school year as Mixed and the achievement gains since opening also as Mixed. This is the only school in this study where the rating given by Edison and The Evaluation Center are the same.