The Effect of Singapore Mathematics on Student Proficiency in a Massachusetts School District: a Longitudinal Statistical Examination

A report by the
Gabriella and Paul Rosenbaum Foundation

October 2009

Gabriella and Paul Rosenbaum Foundation
764 Mt. Pleasant Road
Bryn Mawr, PA 19010-1847
(ph) 610-527-0195 (fax) 610-527-9308
http://www.rosenbaum-foundation.org

Authors:

Madge R. Goldman
President, Gabriella & Paul Rosenbaum Foundation
Vladimir Retakh
Professor of Mathematics, Rutgers University, New Brunswick, NJ
Robert A. Rubin
Resident Scholar, Mathematics Department, Whittier College
Harvey A. Minnigh
President, Earthworks, Inc.

Mathematics and Science Advisory Board
Gabriella and Paul Rosenbaum Foundation

Table of Contents

List of Figures

List of Tables

Summary

We show that for levels of Cumulative Singapore Mathematics experience (CSM) over the period 2002-09, students' scores increase with increasing CSM. For example, North Middlesex Regional School District (NMRSD) students who were 4th graders in 2006 scored much better in 2008 as 6th graders, relative to the overall Massachusetts (MA) student results that we use throughout for comparisons, and this despite the marked improvement over that period in MA student results.[1] Results for all grades 3 through 8 are included at the Singapore Mathematics Effects Between MA and NMRSD and the Singapore Math Effects Within NMRSD sections.

Multivariate and Repeated-Measures analyses strongly indicate a net positive effect of CSM on Massachusetts's Comprehensive Assessment System (MCAS) scores, both overall and by mathematics performance level; societal sub-groups are considered as well. These results are shown in Aggregate and Parallel Data beginning on page 17 and Multivariate Methods beginning on page 36. We speculate that performance gains due to Singapore math take time to develop, since the duration of participation in Singapore math classes has a greater positive impact on test score gains than participation at any given grade level or no participation at all. The most telling argument is at Table 18 on page 43 where we can clearly see the effect of CSM.

Preface

The mathematics and science performance of students has been an increasing source of anxiety in America for some years. This unease has been reinforced by the Trends in International Mathematics and Science Studies (TIMSS: 1995, 1999, 2003) since they all showed U.S. students placing, internationally, within the "average" category; far below the "high performing" category dominated by Asian countries.[1] The 1995 and later TIMSS examination of participating countries' mathematics curricula revealed that the high performing countries all shared a common, focused and coherent scope and sequence in their curricula. Since in this category only Singapore has English for its language of instruction[2], Singapore's math curriculum and its textbooks received a good deal of interest here.

Eventually this interest led to the 2005 study funded by the U.S. Department of Education: "*What the United States Can Learn from Singapore's World-Class Mathematics System (and what Singapore can learn from the United States)*". [2][3] This

---

[1] The 1995 and 2003 TIMSS ranked Singaporean 5th and 8th graders first in the world for math. Top TIMSS rankings: 1995, Singapore, South Korea, Japan, Hong Kong. 2003, Singapore, South Korea, Hong Kong, Taiwan.
[2] Singapore's primary languages are English, Standard Mandarin (or Standard Chinese), Malay and Tamil.
[3] http://www.air.org/news/documents/Singapore%20Report%20(Bookmark%20Version).pdf

report presented comparisons of Singapore's *Primary Mathematics (PM)* textbook series with those of Scott Foresman and of *Everyday Mathematics*. The report also included a comparison of Singapore's 6th grade assessment (Primary School Leaving Examination, PSLE) items with 8th grade items from NAEP[4] and from selected state tests, as well as Praxis[5], again finding in Singapore's favor.

Meanwhile, largely through word of mouth, the hundreds of U.S. home schoolers using Singapore's *Primary Mathematics* textbooks have grown into thousands, but schools have been slower to adopt them. It is only within the last two years that a few school districts as well as a number of individual schools have begun implementing what is colloquially known as Singapore math (SM). But these implementations are recent, and cannot respond to the frequently heard question: *These books do well in Singapore, but how do we know they will do well in the U.S.?* It is to answer that question that the Gabriella and Paul Rosenbaum Foundation has undertaken this longitudinal statistical study of the North Middlesex Regional School District (NMRSD) in Massachusetts, the single school district whose Singapore math implementation is of long standing, having begun in the year 2000.[6]

Acknowledgements

This study would not have been possible without the leadership and support of North Middlesex Regional School District's Superintendent, Dr. Maureen Marshall, who granted us access to the District's data. We also are indebted to Assistant Superintendent Dr. Deborah Brady for providing invaluable help and devoting so much of her time to the project. Special thanks are due to Dr. Mary M. Waight, Associate Superintendent (retired) for summarizing the District history with Singapore Math.

[4] http://en.wikipedia.org/wiki/National_Assessment_of_Educational_Progress

[5] Praxis, one of a series of teacher certification exams written and administered by the Educational Testing Service (ETS). Various Praxis tests are usually required before, during, and after teacher training courses in the U.S. To be certified to teach in most U.S. states, one must earn qualifying scores on all parts of the Praxis I and certain Praxis II Content Area Assessments.

[6] See, in particular Waight, M.M., *The Implementation of Singapore Mathematics in a Regional School District in Massachusetts:2000-2006*, in *Remarks to a National Mathematics Advisory Panel*. 2006: Cambridge, MA.[14]

Singapore Mathematics Implementation at NMRSD

North Middlesex Regional School District (NMRSD) is a rural school district near the border between Massachusetts and New Hampshire, serving the towns of Ashby, Pepperell and Townsend. In response to poor student performance on state mathematics assessments, this district introduced a number of their teachers to the Singapore mathematics (SM) syllabus during a 2000 summer institute for pilot implementation that fall.

For effective implementation, new K-8 school curricula are most easily begun with K-1 or K-2, with another grade added each successive year. However, worried about their entering high school students' inadequate math knowledge, NMRSD chose to address these concerns with an SM pilot in their "feeder" middle schools (grades 5 to 8). The pilot program was quickly extended across classrooms and grades. Kindergarten was added in the 2002-03 year and, as can be seen in Table 1 on page 10, every classroom in grades 1-6 was using the SM curriculum by 2005-06. From this point on, Singapore math was established as the District's official curriculum.

At the same time that NMRSD was establishing the SM curriculum, Massachusetts was implementing higher standards statewide. It was crafting new mathematics standards (based in part on Singapore's) and toughening state tests as well as teacher certification requirements (most recently in 2009). Its schools have responded well and, overall, the State has had considerable success. Participating in TIMSS 2007 as a separate State, Massachusetts earned 5th place overall, solidly ahead of most countries including the U.S. itself. Altogether, Massachusetts was providing a competitive base over the course of NMRSD's Singapore math implementation.

SM Classroom Implementation

Since this study is focused on student outcomes, we do not attempt more than a brief outline of the mechanics of NMRSD's SM implementation.

The SM pilot began in the 2000-01 school year with the use of Singapore's *Primary Mathematics, 3rd Edition*[7] for grades 1-6 and the *New Elementary Mathematics (Syllabus D)* series for grades 7-8.

During the 2001-04 pilot years, students were assigned to SM teachers and non-SM teachers, respectively, by stratified random assignment. The goal was to achieve similar classroom enrollments, gender balances, and comparable proportions of special education and Title I students in each class.

[7] Giving teachers added textbook options, the U.S. Edition of *Primary Mathematics* (pub. 2003) supplemented metric weights and measures with their US counterparts.

Since SM instruction is cumulative, each grade builds on the preceding grades both in SM content and the special SM pedagogy including the "model method" for solving problems, which serves as a bridge to Algebra[8]. For students new to SM, teachers were responsible for providing support for any needed math "catch-up" as well as their own grade's math topics and pedagogy. During this pilot phase of SM implementation, teachers prepared special assessments twice a year that would identify for more attention those topic areas where student understanding seemed weak. Later, in 2007-08, common math assessments for grades 8-10 were implemented followed in 2008-09 by assessments for grades 1-6.

Setting the most effective curricular pacing was and remains a challenge, in part because the SM curriculum sequence is not fully aligned with the Massachusetts Standards. In order to cover topics examined for in the MCAS that are outside of the regular SM learning sequence, teachers developed and followed curriculum maps, usually looking to their pre-SM textbook material for supplementing SM.[9]

Initial teacher preparation for the SM pilot was arranged with nearby Worcester State College. Dr. Richard Bisk[10] gave NMRSD's middle school teachers an intensive 10-day summer institute that was focused on the SM syllabus (math content) linked with the special SM pedagogy.

Part of NMRSD's plan for the institute was to develop their best teachers into "teacher leaders" who would then be designated to help district teachers polish their class-room SM-teaching.[11] Three such teacher leaders did in fact emerge and for some years continued to play a large part in the district's SM implementation.[12] With two additional teacher volunteers, these teachers taught the initial 6 SM classes.

Annual summer teacher workshops have continued to be the norm. In 2001, Dr. Bisk gave a 5-day SM program for teachers of grades 1-4, followed by a separate SM program for teachers of grades 4-10. The 2002 arrangement was somewhat different: two half-day meetings in the Spring, then 5 days in early August and a 6th day meeting later in August. Other institutions covered professional development for NMRSD during 2003 to 2007 and, most recently in 2008 and 2009, Dr. Andrew Chen of Edutron gave an "intensive immersion institute" (boot camp) with SM components for NMRSD middle and high school teachers. Now during the 2009-10 school-year itself, math courses for grades 3-8 teachers are being offered.

---

[8] See *The Model Method*, Singapore Ministry of Education (2009).
[9] Like Massachusetts, all state tests examine mathematics in alignment with their particular curricular standards. The Standards Edition for K-5 Primary Mathematics is fully aligned with California's syllabus, which aligns then closely with MA standards. However, this edition was published only in 2008.
[10] Dr. Richard Bisk, professor of mathematics, has become nationally recognized as an expert on Singapore mathematics.
[11] Math coaches, a separate instructional position, give on-site guidance and help to a school's math teachers. Since school budgets frequently do not allow for this extra position, a common administrative practice is to designate "lead teachers" to perform similar tasks without extra cost to the school.
[12] These teacher leaders also became (and remained for some years) facilitators in Dr. Bisk's professional development classes.

After-school help for struggling students is known to be useful, but schools' ability to offer it depends on available funds. NMRSD's operational budget had only modest sums for such after-school and summer instruction but when such help is available, NMRSD notifies both those students and their parents by letter.

For 2000-04, each elementary school had $2000 for after-school MCAS study, 2 days a week. Each middle school had $3000 for before-school MCAS classes, for 2 days a week for 4 weeks. To this amount, state funds added $100 per student per year, but in 2005, state contribution was decreased to $20 per student per year. Additionally, for under-performing high school students $15,000 was allotted in 2006 and increased to $20,000 in 2007.

Later, more targeted development was also put in place. In 2008, $24,000 was allotted for developing a summer "bridge" math course for incoming 9th graders who had struggled in 8th grade. In 2009-10, double blocks of math have been added for "at risk" students in grades 9 and 10. Middle schools offered an MCAS academy for each grade level for 12 weeks before or after school. In the elementary schools, MCAS after school support was offered for struggling students. However, there continue to be no funds for extra positions of math specialist.

Source/Verification of the Data Employed by the Study

The Singapore mathematics textbook series used by NMRSD schools differ from textbooks used in all other Massachusetts (MA) school districts. The State's student population at each grade is about 70,000 students, while NMRSD's is about 300 students. Thus, a reasonable approach in assessment of the effect of Singapore Mathematics (SM) on NMRSD students' is comparison of their mathematical performance with that of all of MA students. And since SM implementation took place over several years, it was also possible to compare SM versus non-SM student performance within the District itself, as well as with the state as a whole.

The Massachusetts Comprehensive Assessment System (MCAS) which evaluates student, school, and district mathematics performance is an ideal assessment tool. MCAS tests for mathematics are given in grades 3to 8 and 10.[13]

The Massachusetts Department of Education website contains and provides public access to an impressive amount of MCAS as well as demographic and sociological data at the overall state, district and school levels. Yearly MCAS Technical Reports are also issued, especially valuable for the researcher. However, individual

[13] During school years 1998-1999 and 1999-2000, MCAS tested only grades 4, 8 and 10; 2000-01 to 2004-05, grade 6 tests were added; only since 2005-06 have grades 3 to 8 and 10 been tested.

student data are sequestered from public access. It is for district use. The MA department of education has contracted with dataMetrics Software,® Inc. to make specialized software programs (the TestWiz®) available to interface the MCAS with the District's own database. For example with TestWiz, teachers can call up students' MCAS results together with their attendance record for the current and/or past years.

Once given District permission to access individual student data, TestWiz enabled us to track individual students and their academic performance in mathematics during their years in District schools. However, at this point we discovered a fundamental shortcoming in the Massachusetts record-keeping system. It does not record the particular curriculum used in any subject. While math scores are available, the data collection system does not have the capability to record the textbook series in use. Thus, determining a student's Cumulative SM (CSM) exposure was a challenge.

Among the resources we used to track down the CSM of individual students, a table included in Dr. Mary Waight's testimony to the National Mathematics Panel was a useful guide. This illustrated the by-class and by-grade progression of Singapore mathematics implementation at NMRSD. We adapt a portion as Table 1, below.

Table 1. Number of Singapore Math Classes at NMRSD

	Kindergarten	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
2000-01	**0**	0	0	0	**0**	1 of 17	1 of 17	2 of 16	2 of 18
2001-02	**0**	2 of 17	0	0	2 of 17	5 of 17	3 of 17	4 of 16	3 of 18
2002-03	0	6 of 17	10 of 16	3 of 16	7 of 17	10 of 17	7 of 17	5 of 16	5 of 18
2003-04	0	9 of 17	12 of 16	14 of 16	10 of 17	11 of 17	9 of 17	7 of 16	5 of 18
2004-05	3	14 of 17	13 of 16	16	11 of 17	17	12 of 17	8 of 16	6 of 18
2005-06	3	17	16	16	17	17	17	12 of 16	14 of 18
2006-2007	3	17	16	16	17	17	17	12 of 16	14 of 18
2007-2008	3	17	16	16	17	17	17	16	18

Although Massachusetts electronic record-keeping of individual student MCAS scores was only put in place in 2003, NMRSD kept its own meticulous records for students in the SM pilot classes. This enabled us to obtain data for the 2001-02 year for SM classes, such as the MCAS scores for the pilot's 4th graders. (NMRSD did not similarly keep individual MCAS data of their 2000-2002 non-SM classes.)

SM implementation extended as shown above in Table 1. Each grade-year with 100% SM classes adds a year to students' CSM. However, determining a student's CSM for grade-years when SM-participation was less than 100% was tricky. For each such grade year, Testwiz® extracted a list of all students in each instructor's class from the District databank. We then identified (from paper school records) all SM-teachers in each grade-year. Then, and only then, students (identified by their unique school IDs) who were in known SM-teacher classes could be assigned to SM experience for that grade-year[14].

This roundabout method does have an advantage: it allowed us to find and count students' SM experience even for grade-years when no MCAS is given. A sample concatenated file for a student set counting 5th grade SM experience in this way (no MCAS for 5th grade in that year) is shown at Table 2.

Table 2. Sample Cumulative SM data file.

Student	4th grade Raw MCAS Score	4th grade SM	5th grade SM	5th grade CSM	6th grade Raw MCAS Score	6th grade SM	6th grade CSM
#1	40	Yes	Yes	2	45	Yes	3
#2	49	Yes	Yes	2	50	Yes	3
#3	52	Yes	Yes	2	48	Yes	3
#4	33	No	Yes	1	43	Yes	2
#5	39	No	Yes	1	46	Yes	2
#6	22	No	Yes	1	23	Yes	2
#7	50	No	Yes	1	48	Yes	2
#8	47	Yes	Yes	2	48	Yes	3

Cohort – CSM data

We began evaluation of the effect of SM on NMRSD students' performance over time by choosing groups of students meeting three conditions.

(1) All of their test scores in Massachusetts's Comprehensive Assessment System (the MCAS) for 2002 through 2008 were available.

[14] District teacher-student records were however incomplete: this data was missing for 2003-04 and 2007-08.

(2) Their cumulative Singapore mathematics experience (CSM) over 3 or more years could be verified.
(3) All students within the group were in the same grade and school-year.

We call these groups of students cohorts or vectors. For example, the vector of the students who were in the 4th grade in 2003-04 would be followed year by year through to 8th grade in 2007-08.

For analysis we chose four such cohorts, V1 to V4, with over 1,000 student-trials[15] in each cohort. This provides for a very powerful study.

Note that MCAS results are not available for all grade years. Until 2001-02, only grades 4, 8 and 10 were tested; from 2002-03 on, grade 6 was also tested. Finally in 2005-06 MCAS began testing all grades 3 to 8. Given the limited MCAS testing in years before 2005-06, the four cohorts we chose have the most complete data. Vectors are shown in Table 3 on page 14. Note, also:

1) While no MCAS was administered in 5th grade in 2001-02 nor in 7th grade in 2004-05, we do have cohort V1's SM exposure data for these years from NMRSD sources.
2) While 10th grade results were available, the study did not include these in most of the analyses because SM was not used after 8th grade. A number of graphs do include 10th grade results, for completeness; see, for example, Figure 13 on page 49.

Parallel Groupings

Using the extensive NMRSD data to full advantage, we also examine for the relative math performance over classes, for example of students in 4th grade in 2004, in 2005, in 2006 and so forth. That is, these are different classes, in the same grade for different years; we call these groupings *parallels.*

Categorical Data

We also examine MCAS results that are derived from the overall grade-year means[16]; measures of dispersion and reliability are reported in four categories (performance levels) for ranges of raw scores. This sort of categorization is used in Massachusetts and most other states: for MA the categories are "Advanced, Proficient, Needs Improvement, Warning/Failing". These categories provide easy iden-

[15] For example, cohort V4 was tested five times during grades 4 through 8. V4 totaled 1,711 student-trials, ranging from 357 student-trials in 2005 to 345 in 2009.
[16] Both raw and scaled mean scores are provided in state reports. Scaled mean scores are described in their *Technical Reports*, but were not used in this study.

tification of performance level, and allow for comparisons between individual student performance over time, and for groups by grade-year. Since MCAS provides the number of students in each of these categories for 1999-2009, categorical analyses were used over these years to compare MA and NMRSD performance by school-year, by grade-year, and by CSM exposure.

A later section that most directly addresses the effects of CSM exposure, Multivariate Methods, begins on page 36.

Methods

The MA and NMRSD test results are expressed as "mean scores". The scores themselves report the number of correctly answered test problems. Thus for grades 4 to 9, the maximum MCAS score is 54. However, in years when the maximum MCAS score differs from 54, we use "adjusted means" (or "standardized" or "scaled means"). For grade 3, where the maximum score is 40, scores are adjusted by the ratio 54/40. For most of this report Grade 10 is not reported; but grade 10's scores would be adjusted by 60/54.

When looking at results, it may be helpful for the non-statistician to see mean scores also in the form of percents. We call these "coded" means; a student score of 39 out of 54 MCAS problems would have 72% of the problems correct.

As described below, comparisons were carried out on both mean (or continuous) data and on categorical data. Both univariate and multivariate statistical methods were used, on MA versus NMRSD results and on within-NMRSD results. We also examined the effect on math results of sub-categories representing gender; generalized economic background of students; and whether students were enrolled in special education. In examining each year's mathematics performance, the study sought to determine whether continuing in SM over several years – that is, Cumulative Singapore Math (CSM) exposure – makes a difference. For this we relied on Multivariate, Repeated-Measure methods.

Student Outcomes by CSM

Continuous data

MA-MCAS results were gathered mostly from MCAS's annual *Technical Reports* [3-13]. MCAS data are reported for each tested grade in each school, in each District, and for the State as a whole. By NMRSD permission, MCAS data for their individual students (NMRSD-MCAS) was accessed via the TestWiz® facility.

In the Singapore Mathematics Effects Between MA and NMRSD section, beginning on page 22, analyses consisted of one-way analysis of variance (ANOVA), t-tests, means tests within NMRSD for continuous, or mean, data for the cohorts used in this study.
Multivariate measures also use cohorts or vectors of students over years at NMRSD and compare the cohorts' MCAS test results (mean scores) with the State's.

Table 3. Vectors are shown moving through grade-years. White cells represent grade-years with less than 100% of the classes receiving Singapore math instruction.

YEAR	*Grade*							
Cells in this color are 100% SM	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
2000								
2001								
2002								
2003								
2004								
2005								
2006								
2007								
2008								
2009								

V1
V2
V3
V4

Cohort and Parallel Data

Using NMRSD's extensive data to good advantage, we also examine relative performance over classes, for example of students in 4th grade in 2004, 2005, 2006 and so forth. These are different classes in the same grade for different years; we call these groupings parallels.

While the gaps in teacher-student information handicap a fuller investigation of some of our cohorts' relationship between CSM and math performance, we can use the parallel vectors to compare student performance by relative experience in SM instruction. For example, we have MCAS test data for 4th grade from 1998 on. If we look at mean student performance in 2000-01, 2001-02, 2002-03, 2003-04 and thereafter (grouped), we may draw inferences from the differences in student performance that the parallel vectors highlight. While not individually related to student CSM, relationships will be suggestive of CSM exposure for the classes as a whole.

Student Outcomes by Teacher

In addition to the curriculum used: a major, perhaps *the* major, determinant of student performance - besides aptitude and application - is the experience of the teacher in that curriculum. That effect may be expressed in various ways, but we chose to use student results by individual teachers' years of SM experience. Student results for individual teachers are shown at Figure 1. It is clear that teacher results vary considerably.

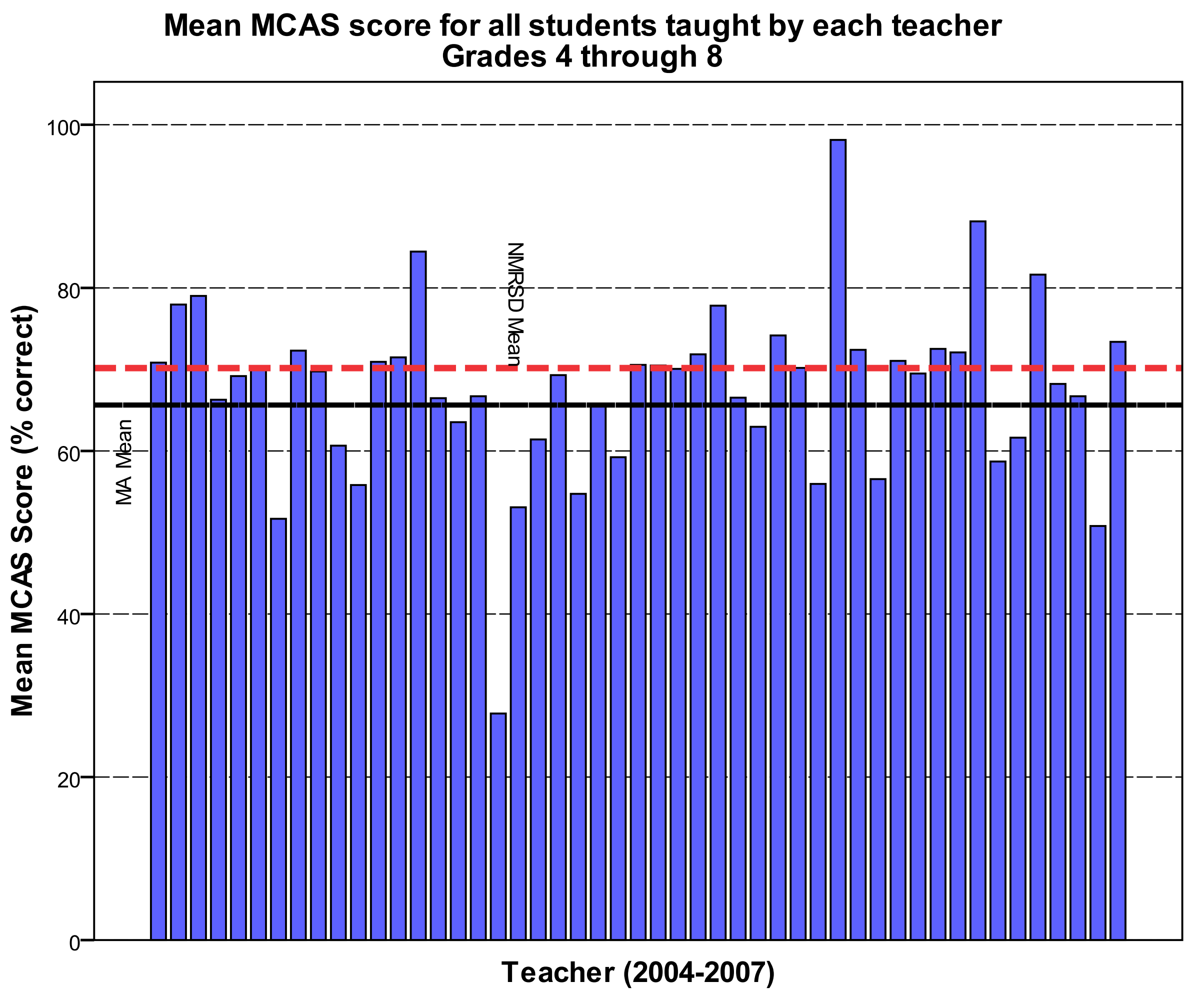


Figure 1. Student Results by Teacher.

We selected teachers who, in 2006, had between 0 and 5 years of experience with SM, and followed their students over parallels consistent with their experience. The measure of performance was mean student MCAS score over the period 2004-2007. These results are shown at Figure 2, below. The heavy, dashed red line indicates the MCAS mean score of NMRSD students overall Note that this is higher than the MA students' mean scores.

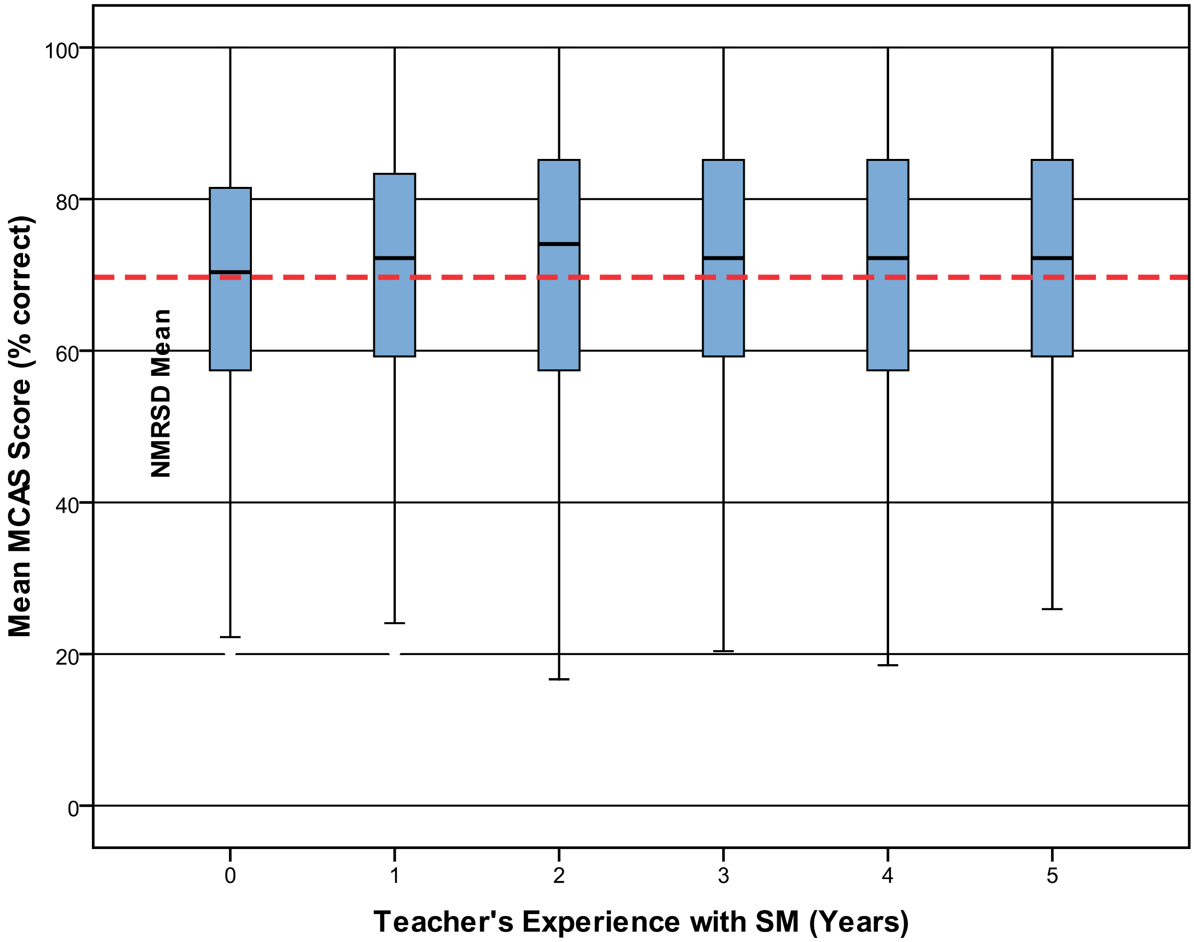


Figure 2. Mean Student MCAS Results by Years of Teacher Experience.

These figures show that there is great variability – often around 50% - in teachers' success with their students, whatever SM experience teachers had. (Notice the "belt" of each figure, their students' mean scores.) Confounding factors are too numerous and ill-understood to allow further clarification. (Note that teacher effects are also considered in the multivariate section of this report.)

Aggregate and Parallel Data

One method of measuring the effect of SM over time is to review student MCAS scores for the same grade over the period of study available to us. Since SM was introduced in some classes and some grades over a period of years, students in later grades who had higher CSM might have different results. These differing results will help quantify the efficacy of SM.

For example, referring to Table 1, (page 10), 8th graders in 2003-04 had 5/18 chances of being in an SM class and had 5/16 chances in 7th grade, 3/17 chances in 6^{th} grade and so forth. 8^{th} graders in 2004-05 had approximately 6/18, 7/17, 7/17, 5/17 and 0 chances of being taught SM since 2001; while 8^{th} graders in 2005-06 had a greater likelihood (14/18, 8/16, 9/17, 10/17 and 2/17) of exposure to SM. MCAS scores may reflect the effects of SM exposure.

Figure 3 below shows these SM results overall for the different grade-years, while Table 4, beginning on page 19, shows individual SM student results. The trends are (with the exception of 4^{th} grade) positive over the study period. Comparing Grades 7 and 10 over years by ANOVA, we find they are significantly different (Levene's Statistic 1.267, P=0.284 and F=2.896, df 3,1389 and P=0.034 for 7^{th} grade and Levene's 0.758, P=0.580 and F=39.441, df 5,1780 and P=0.000 for 10^{th} grade). For the other grades, all Levene's Statistics suggest significant differences in variance.

The 4^{th} grade results are somewhat puzzling, but nevertheless follow the MA trend (see Figure 12 on page 48). Note that the same SM students who score poorly in 4^{th} grade in 2003-04 score higher than MA in 6^{th} grade in 2005-06, and that trend continues.

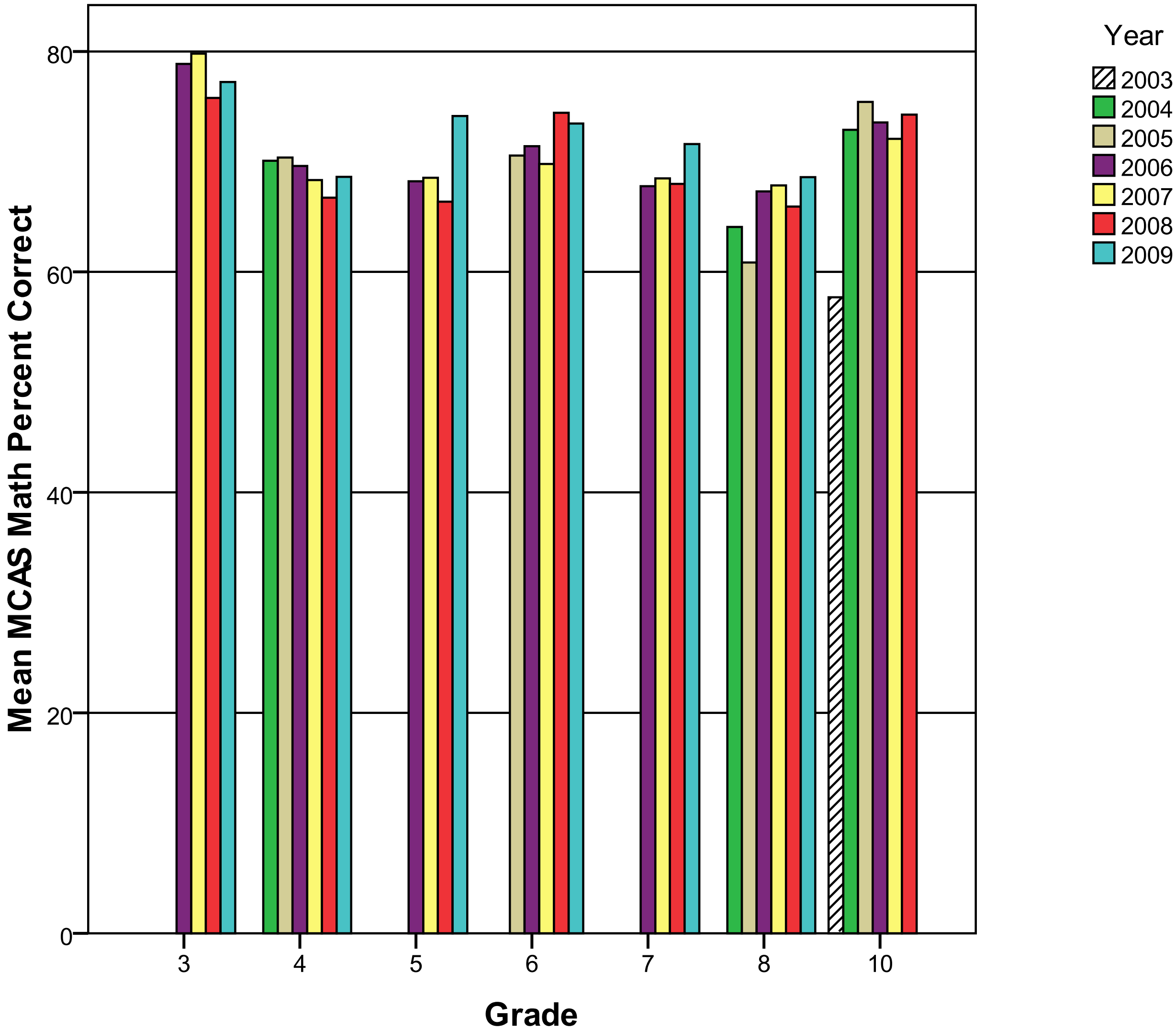


Figure 3. Students' Mean Percent Correct by Grade Parallels, 2003-2009

Table 4. SM Students' Mean Scores by Category and Grade 2003-2009

Year	Category	Grade Mean (percent correct)						
		3	4	5	6	7	8	10
2003	MATH Raw							57.70
	Number Sense							61.57
	Patterns, Relationships & Algebra							57.44
	Geometry							55.55

Table 4. SM Students' Mean Scores by Category and Grade 2003-2009

Year	Category	Grade Mean (percent correct)						
		3	4	5	6	7	8	10
	Measurement							68.33
	Data, Statistics & Probability							47.95
2004	MATH Raw		70.1				64.09	72.89
	Number Sense		70.8				63.97	79.25
	Patterns, Relationships & Algebra		71.4				66.04	72.03
	Geometry		67.5				64.04	71.39
	Measurement		57.4				51.36	63.79
	Data, Statistics & Probability		78.6				75.34	74.47
2005	MATH Raw		70.4		70.56		60.86	75.42
	Number Sense		73.0		76.24		56.40	72.03
	Patterns, Relationships & Algebra		73.7		74.54		69.27	70.31
	Geometry		63.7		63.17		55.46	79.91
	Measurement		64.3		59.87		45.21	77.83
	Data, Statistics & Probability		71.7		68.71		70.88	80.01
2006	MATH Raw	78.87	69.6	68.23	71.41	67.78	67.30	73.57
	Number Sense	80.15	70.4	71.05	75.36	67.85	63.45	79.75
	Patterns, Relationships & Algebra	83.69	78.0	73.29	76.98	66.05	68.78	70.23
	Geometry	69.52	51.5	59.62	55.75	73.54	63.06	71.46
	Measurement	74.87	58.2	64.68	64.04	61.19	66.33	72.61
	Data, Statistics & Probability	81.44	77.2	63.65	73.13	70.76	73.65	73.49
2007	MATH Raw	79.81	68.3	68.54	69.80	68.49	67.84	72.07
	Number Sense	75.64	70.3	70.73	72.37	61.42	68.06	73.44
	Patterns, Relationships & Algebra	79.18	72.2	69.90	74.16	71.45	70.40	66.40
	Geometry	80.96	58.6	63.62	60.14	75.32	69.83	72.93
	Measurement	82.39	62.0	55.40	77.68	56.86	60.37	78.38
	Data, Statistics & Probability	84.61	70.9	78.00	58.21	76.62	67.75	73.53
2008	MATH Raw	75.78	66.7	66.37	74.43	67.98	65.92	74.27
	Number Sense	73.64	65.5	68.44	77.77	67.92	67.46	73.51
	Patterns, Relationships & Algebra	82.10	79.0	67.72	78.43	73.96	66.44	76.60
	Geometry	68.68	61.4	69.41	75.45	61.25	64.61	74.00
	Measurement	76.58	60.1	55.41	61.92	58.70	56.72	67.89

Table 4. SM Students' Mean Scores by Category and Grade 2003-2009

Year	Category	Grade Mean (percent correct)						
		3	4	5	6	7	8	10
	Data, Statistics & Probability	76.45	64.0	66.42	70.26	69.91	70.17	77.39
2009	MATH Raw	77.24	68.6	74.14	73.46	71.60	68.59	
	Number Sense	78.30	69.6	74.31	75.14	70.59	73.61	
	Patterns, Relationships & Algebra	79.03	66.3	82.35	78.71	74.16	69.28	
	Geometry	65.85	73.4	58.31	69.99	70.54	62.45	
	Measurement	71.95	62.0	72.16	66.34	65.33	68.73	
	Data, Statistics & Probability	83.29	70.0	75.24	70.21	74.25	65.54	

Results for regressions of CSM against MCAS Raw Mathematics Score each grade-year for which we have test results are at Table 5, below. Please recall that only a very small portion of the variability in scores is accounted for by the regression – generally about 5-10 percent. Two of the grade-years for which we have a good deal of data are not significantly different; these are 4th and 6th grades.

Table 5. Regression for Grade by Year

	Coefficient	t	P
Grade 3	-.061	-2.186	.029
Grade 4	.018	.884	.377
Grade 5	.081	3.300	.001
Grade 6	.031	1.514	.130
Grade 7	.057	2.385	.017
Grade 8	.080	4.129	.000
Grade 10	.202	8.724	.000

Results

Univariate Measures

We looked at Singapore Math effects between MA and NMRSD student' results and also within NMRSD's SM and non-SM students.

Singapore Mathematics Effects Between MA and NMRSD

Continuous Data

We first examined MCAS scores for NMRSD and for MA. Standardized grade means for MA and for NMRSD are compared in Figure 4, for the years 2003-2008. Note that on a grade basis, NMRSD scores are consistently higher, occasionally significantly so, for example, 3rd grade.

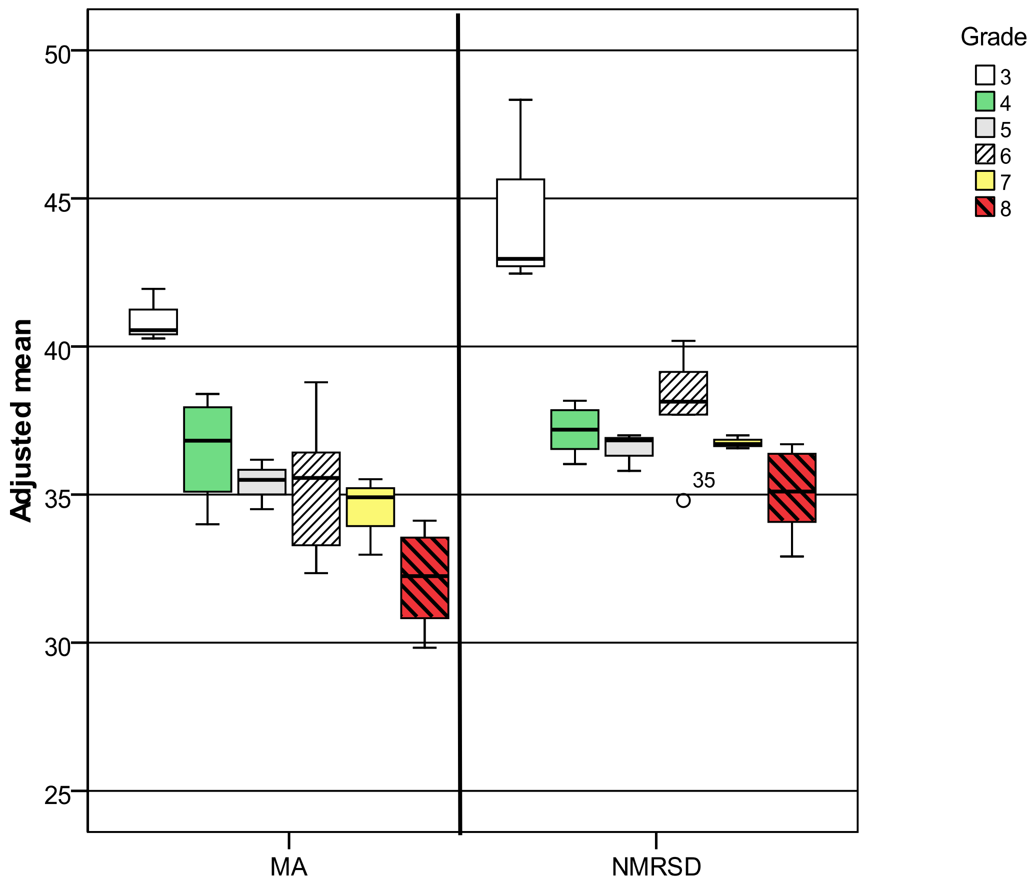


Figure 4. MA versus NMRSD mean MCAS scores, by grade, averaged for years 2002-03 to 2007-08.

Other ways to view the same data are provided in Table 6 and in Table 9; there, the standardized mean and a coded, standardized mean are shown, respectively. Table 6 provides an additional measure: the difference between NMRSD and MA coded scores (i.e., NMRSD coded score minus the MA coded score). In all but three (of 33) grade-years, NMRSD scores were higher than MA scores. Figure 5 on page 24 shows this data graphically. Grade 10 is included for completeness.

Table 6. Adjusted Mean and Coded Mean MCAS scores by grade by year

Year	Grade	MA Adjusted Mean	MA Mean (coded 54=100)	NMRSD Adjusted Mean	NMRSD (coded 54=100)	NMRSD minus MA
2003	4	34	63	37	68	5
	6	32	60	35	64	4
	8	30	55	34	63	8
	10	29	54	31	58	4
2004	4	35	65	38	70	5
	6	36	66	39	72	6
	8	31	58	35	64	6
	10	35	65	39	73	8
2005	4	37	68	38	71	3
	6	33	62	38	70	8
	8	31	57	33	61	4
	10	35	65	41	75	10
2006	3	42	78	42	79	1
	4	38	71	38	70	-1
	5	35	64	37	68	4
	6	36	66	39	71	5
	7	33	61	37	68	7
	8	34	62	36	67	5
	10	35	65	40	74	9
2007	3	40	75	43	80	5
	4	38	70	37	68	-2
	5	36	67	37	69	2
	6	36	67	38	70	3
	7	35	65	37	69	4
	8	33	62	37	68	6
	10	36	67	39	72	5
2008	3	41	75	48	90	15
	4	37	68	36	67	-1
	5	36	66	36	66	0

Table 6. Adjusted Mean and Coded Mean MCAS scores by grade by year						
Year	Grade	MA Adjusted Mean	MA Mean (coded 54=100)	NMRSD Adjusted Mean	NMRSD (coded 54=100)	NMRSD minus MA
	6	39	72	40	74	2
	7	36	66	37	68	2
	8	34	63	36	66	3
	10	36	66	40	73	7

An ANOVA on these mean scores is significant (F=11.406, P=0.001, df=1,64) and this shows that, overall, NMRSD mean MCAS scores are significantly higher than those of Massachusetts for all grade-years for the 2003-2008 period.

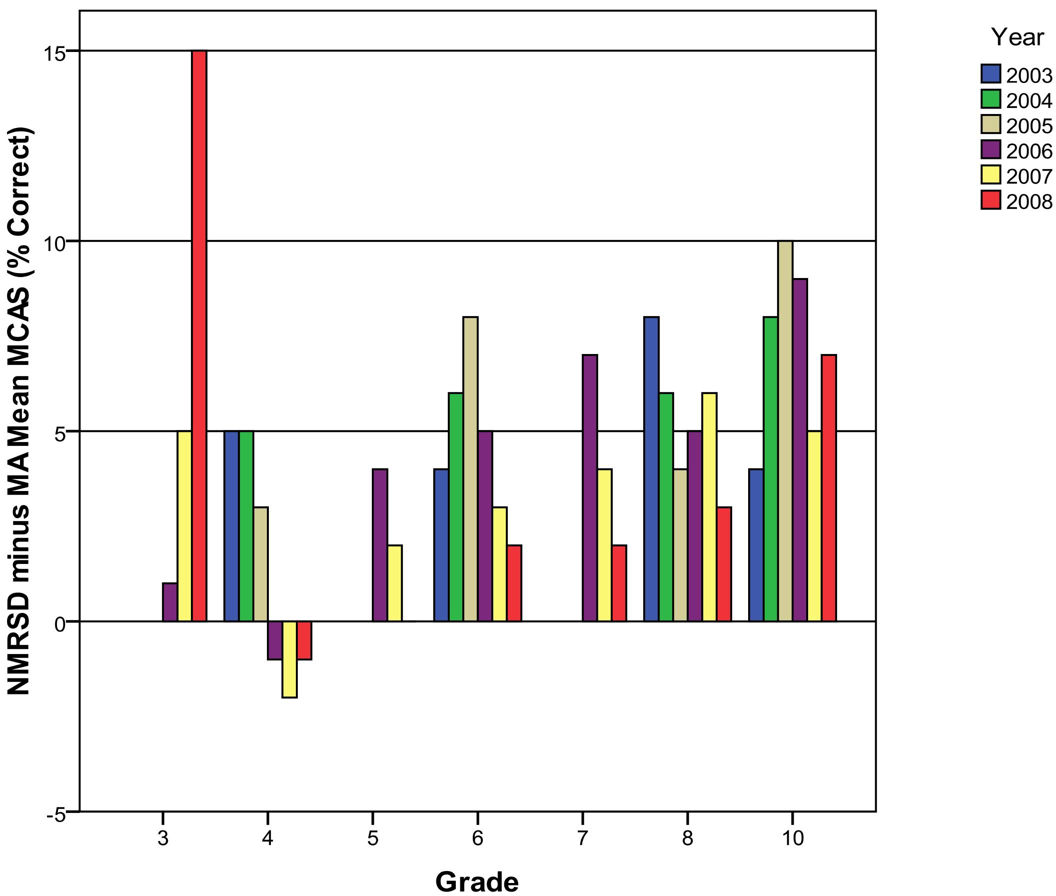


Figure 5. Differences between NMRSD and MA Students' MCAS Coded Scores, 2002 to 2008

For the period 2003-2008, it is clear that NMRSD scores on the MCAS are regularly higher than MA's. A number of trends are apparent. As Figure 12 and Figure 13 (pages 48 and 49, respectively) also show, MA and NMRSD scores do seem to follow the same general upward trends: with 4th grade scores the only exception.

While regression on the two sets of data pairs (MA mean and NMRSD mean) is significant (Adjusted $r^2 = 0.626$, $F=54.626$, $P < 0.001$) this only reflects the trends we can see and does not connect them to a cause. Now we will pass to testing the contribution of Singapore Math to these results.

Singapore Math Effects Within NMRSD

For Singapore Math's effects, we rely on the Vectors described in Table 3 on page 14 and in Cohort – CSM data beginning on page 11. Given the well-known complexity of factors bearing on student performance in testing, we could not expect to demonstrate significant differences in mean scores based just on univariate statistics. However, there are some suggestive findings, including some of those that led us to this study. Thus, it is useful to present our univariate results. For all students in the cohorts, Figure 6 below shows the relative frequency of MCAS Mathematics Raw scores (that is, how many students got a particular mean score) with and without Singapore Math experience for all scores in vectors V1 to V4. For the two data sets the medians are 37 and 39, means are 35.58 and 37.33 for CSM=0 and CSM>0, respectively. Student's t is -4.142 (equal variances not assumed (Levene's F=4.151, p=0.042), df 838, p=0.000).

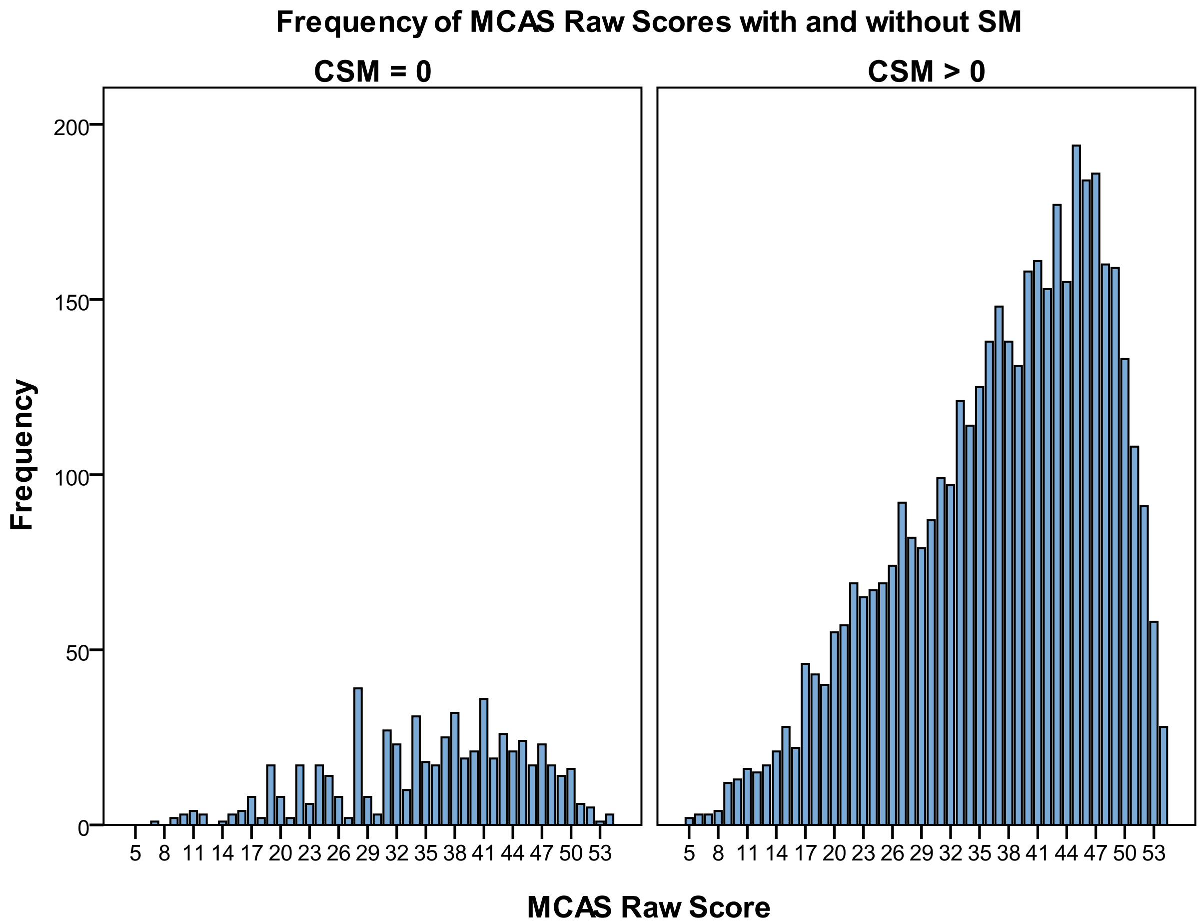


Figure 6. Frequency of MCAS Raw Scores Without and With CSM for cohorts V1 to V4.

Figure 7 on page 28 below shows the effect of CSM on mean MCAS scores by grade. For example, grade 5 mean scores increase from about 35.5 to 37 to 38 for students with CSM's of 0, 1 and 2 years, respectively. Note that the figure reflects the MCAS testing history for grades; since MCAS testing of grade 5 only began in 2006, there are only the 3 years of CSM experiences applicable to 5th grade MCAS scores in those academic years, i.e., 2005-2006, 2006-07 and 2007-08.

Correlations

Since CSM is ordinal, a non-parametric measure of correlation, Spearman's rho, was used. For correlations, only complete Cohorts were analyzed; results for subsets of CSM (i.e., a subset of grades of SM experience) were not substantially different. Results for the Cohorts for MCAS scores[17] are shown in Table 20; 4 of 7 are not significantly correlated. Note, however, that all the rho's are positive, i.e., MCAS scores rise with increased CSM exposure.

Correlations between CSM and the Social Parameters considered are presented in Table 21. The Free Lunch category (an indirect measure of family income) and Special Education are always significant. On the other hand, Gender (i.e., male, female) is not usually significant for all data in the Cohorts; the exceptions are in Cohort V1. Nevertheless, gender is significant for some grade-years, although there doesn't appear to be a pattern to this. Social Parameter comparisons are shown in Table 21, below.

One-way ANOVA, t-tests and Regression

It is easy to show general improvement with increasing exposure to Singapore Math. For example, Table 8, below, shows t-tests by grade-year and for all the cohort V1, V2, V3 and V4 students; while Table 9 shows the relationship between CSM and student performance for all students, and for students in the Social Parameter sub-categories. By regression, for all students, these are significantly related (i.e., CSM may predict scores) (F=7.767, df 1,4,648, P=0.005)[18]

MCAS Scores are significantly different (F=7.244, df 1,4,241, P=0.007) for Free Lunch compared to those not receiving Free Lunch. In the same way, differences for Special Ed students they are also significant (F=6.972, df 1,587, P=.009).

While the coded standardized mean scores may not appear different, they are, and CSM is a significant factor. The repeated measures dimension of these analyses are much more compelling. These are treated in the Multivariate Methods section beginning on page 36.

[17] Scores are raw scores for most grade-years and adjusted scores for 3rd grade scores; only Cohorts V1 and V4 are shown in Figure 3.
[18] Regressions are run on adjusted mean scores, while coded adjusted mean scores are shown in Table 9.

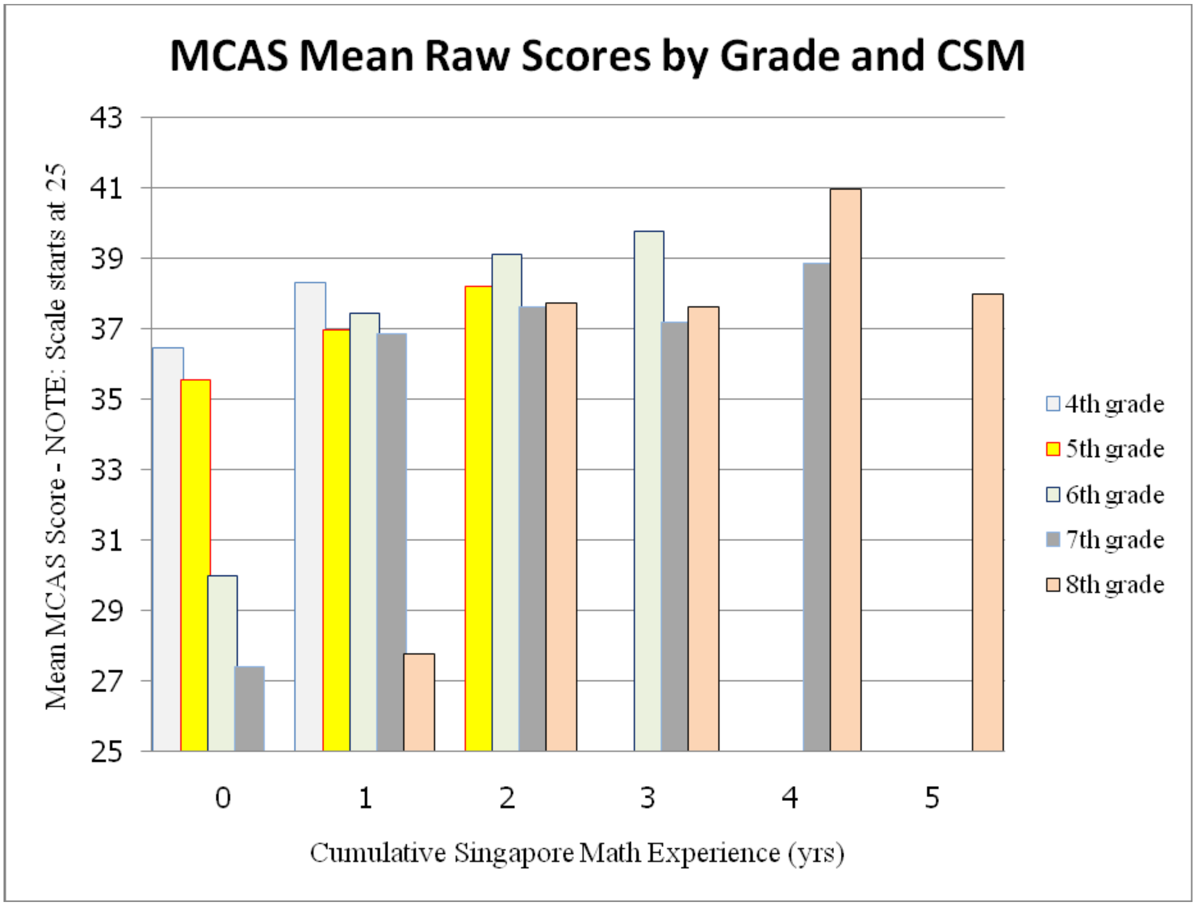


Figure 7. CSM Experience Effect on MCAS Scores

Table 7. MCAS Mean Scores by Test Category – All Data 2002-2009							
MCAS Test Category	SM	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean	
						Lower	Upper
Raw Math	No	640	31.53	9.218	0.364	30.81	32.24
	Yes	9215	32.88	8.986	0.094	32.7	33.06
Number Sense	No	132	10.49	5.364	0.467	9.57	11.42
	Yes	9039	11.43	3.827	0.04	11.35	11.51
Patterns and Relationships	No	132	7.49	2.846	0.248	7	7.98
	Yes	9039	9.48	3.132	0.033	9.42	9.55
Geometry	No	132	3.66	2.301	0.2	3.26	4.06
	Yes	9039	4.45	1.787	0.019	4.42	4.49
Measurement	No	132	3.91	2.572	0.224	3.47	4.35
	Yes	9039	4.26	1.852	0.019	4.22	4.3
Data, Statistics & Probability	No	132	5.76	2.569	0.224	5.32	6.2
	Yes	9039	6.81	2.404	0.025	6.76	6.86

Student's t-tests for all data based on SM are also significant for all mathematical categories except for Measurement.

Table 8. t-tests for MCAS mean score by CSM				
Grade	t	n	P	CSM range
4	102.167	7	0.000	0-1
5	190.3	2	0.000	1-3
6	24.452	11	0.000	0-3
7	26.025	8	0.000	0-4
8	14.952	10	0.000	0-5
all grades, all years	44.981	42	0.000	0-5

Table 9. Coded standardized means for all students, all Cohorts by SM experience					
Coded Adjusted Mean Scores					
Cumulative SM experience	All Students	Not Free Lunch	Free Lunch	Not Special Ed	Special Ed
0	67	68	59	70	49
1	70	71	63	73	48
2	72	73	64	75	49
3	71	72	61	73	57
4	69	70	59	71	53
5[19]	70	71	67	72	61

ANOVA results for individual Cohorts are presented in Table 22 beginning on page 51 for all scores; in Table 23 beginning on page 53, for the Free Lunch categories; and in Table 24, beginning on page 54 for the Special Education categories.

In each of these, CSM categories with very low N's (low number of students) are excluded from analyses. For the raw scores, 1 of 7 comparisons is significant by ANOVA; for Free Lunch, 9 of 24 and for Special Ed, 1 of 11 are significant.

t-tests for all data for the various math sub-categories are shown in Table 10, below. In every case, mean scores for students with SM experience are higher than those with none. Note, also, that there are no more than 640 students with no SM math experience, pointing-up again the major problem with a study conducted on historical data.

[19] Only 143 students (3%) have 5 years of CSM.

Table 10. t-test results for all tested students 2003-2009 for math sub-categories							
	SM	N	Mean	Std. Deviation	t[20]	df	sig.
Math Raw	0	640	31.53	9.218	-3.681	9,853	0
	1	9,215	32.88	8.986			
Number Sense	0	132	10.49	5.364	-2.007	133	0.047
	1	9,039	11.43	3.827			
Pattern Recognition	0	132	7.49	2.846	-7.258	9,169	0
	1	9,039	9.48	3.132			
Geometry	0	132	3.66	2.301	-3.947	133	0
	1	9,039	4.45	1.787			
Measurement	0	132	3.91	2.572	-1.557	133	0.122
	1	9,039	4.26	1.852			
Data, Statistics & Probability	0	132	5.76	2.569	-5.004	9,169	0
	1	9,039	6.81	2.404			

Finally, t-tests on student results by socio-economic categories by SM are shown at Table 11. Again, all results for students with SM experience are higher, with the notable exception of Free Lunch-non-special Education students. A number of these results are significant, including all but 3 categories of Special Education students. Those were Free Lunch-Special Ed students for Math Raw, Pattern Recognition and Data, Statistics and Probability. It does appear from these that SM has generally had a beneficial result on students with disadvantages. It appears, as well, that the task of improving the results obtained by economically disadvantaged students may be somewhat more intractable than for Special Education students.

Table 11. T-tests on Social Categories by SM – All data 2003-2009									
Free Lunch	Special Ed	MCAS score	SM	N	Mean	Std. Deviation	t	df	p
Free	Not SPED	Math Raw	0	36	37.03	8.994	1.212	622	0.226
			1	588	34.95	10.065			
		Number Sense	0	9	13.22	3.993	1.773	586	0.077
			1	579	11.01	3.720			
		Patterns,	0	9	7.78	2.489	-1.220	586	0.223

[20] The negative sign of t-test statistics is an artifact of coding "no SM" as 0 and "SM" as 1.

Table 11. T-tests on Social Categories by SM – All data 2003-2009

Free Lunch	Special Ed	MCAS score	SM	N	Mean	Std. Deviation	t	df	p
		Relations and Algebra	1	579	9.07	3.166			
		Geometry	0	9	3.89	2.205	-0.423	586	0.673
			1	579	4.14	1.785			
		Measurement	0	9	4.89	2.369	1.146	586	0.252
			1	579	4.16	1.879			
		Data, Stat-Probability	0	9	6.78	2.635	0.448	586	0.654
			1	579	6.42	2.397			
	Yes SPED	Math Raw	0	33	22.27	10.220	-0.422	237	0.674
			1	206	23.00	9.024			
		Number Sense	0	16	4.56	2.449	-3.178	217	0.002
			1	203	7.35	3.437			
		Patterns, Relations and Algebra	0	16	5.06	2.380	-1.179	217	0.240
			1	203	5.90	2.765			
		Geometry	0	16	1.06	1.063	-3.935	217	0.000
			1	203	2.77	1.706			
		Measurement	0	16	1.44	1.031	-2.969	217	0.003
			1	203	2.70	1.680			
		Data, Stat-Probability	0	16	3.56	1.861	-0.942	217	0.347
			1	203	4.14	2.405			
Not Free	Not SPED	Math Raw	0	491	37.53	8.553	-2.621	5845	0.009
			1	5356	38.61	8.738			
		Number Sense	0	62	14.82	2.323	9.506	64	0.000
			1	5205	11.98	3.511			
		Patterns, Relations and Algebra	0	62	9.19	1.906	-3.770	64	0.000
			1	5205	10.12	2.909			
		Geometry	0	62	5.18	1.635	2.347	5265	0.019
			1	5205	4.67	1.678			
		Measurement	0	62	5.65	1.976	4.070	62	0.000
			1	5205	4.62	1.709			
		Data, Stat-Probability	0	62	7.32	1.914	0.783	5265	0.434
			1	5205	7.10	2.219			
	Yes SPED	Math Raw	0	72	23.93	8.731	-2.493	882	0.013
			1	812	26.96	9.980			
		Number Sense	0	37	5.89	3.550	-4.346	834	0.000
			1	799	8.67	3.807			
		Patterns,	0	37	5.65	2.508	-2.773	834	0.006

Table 11. T-tests on Social Categories by SM – All data 2003-2009									
Free Lunch	Special Ed	MCAS score	SM	N	Mean	Std. Deviation	t	df	p
		Relations and Algebra	1	799	7.08	3.098			
		Geometry	0	37	2.24	1.786	-3.388	834	0.001
			1	799	3.27	1.808			
		Measurement	0	37	2.11	1.696	-3.049	834	0.002
			1	799	3.02	1.775			
		Data, Stat-Probability	0	37	4.11	1.776	-2.705	42	0.010
			1	799	4.93	2.418			

Categorical Data – Both Between NMRSD and MA and Within NMRSD

Much of the available data on MCAS is in the form of categorical data, where students are assigned a proficiency category based on test results; the categories were described earlier. The great attraction is that we have data from 1998-2008, eleven years. These data for the period 1998-2008 are shown in Figure 8 where the difference (%NMRSD students in category minus %MA students in category) is plotted. It is clear that NMRSD students normally perform somewhat better than their statewide contemporaries. In Figure 8, while NMRSD's Advanced category is higher than that of MA in 3 grades, tied in one and lower in 3 other grades, NMRSD always has a higher proportion of students than does MA in the Proficient category. Likewise, NMRSD always has fewer students in the Failing category.

A word on interpretation: for the category differences, there are two "good" results. One, when NMRSD minus MA is positive, i.e., a higher proportion of NMRSD students are in the Advanced or Proficient categories. The other is at the other end of the spectrum, when NMRSD minus MA is negative, i.e., a lower proportion of NMRSD students are in the Needs Improvement or Failing categories. For this reason, each of the category graphs below are presented for all categories and also for the categories grouped as above, i.e., [Advanced and Proficient] and [Needs Improvement and Warning/Failing]. An example of this format is shown in Figure 10 on page 35. Again, we have included 10^{th} grade in this chart for completeness, since before 2006 MCAS tested only grades 4, 6, 8 and 10.

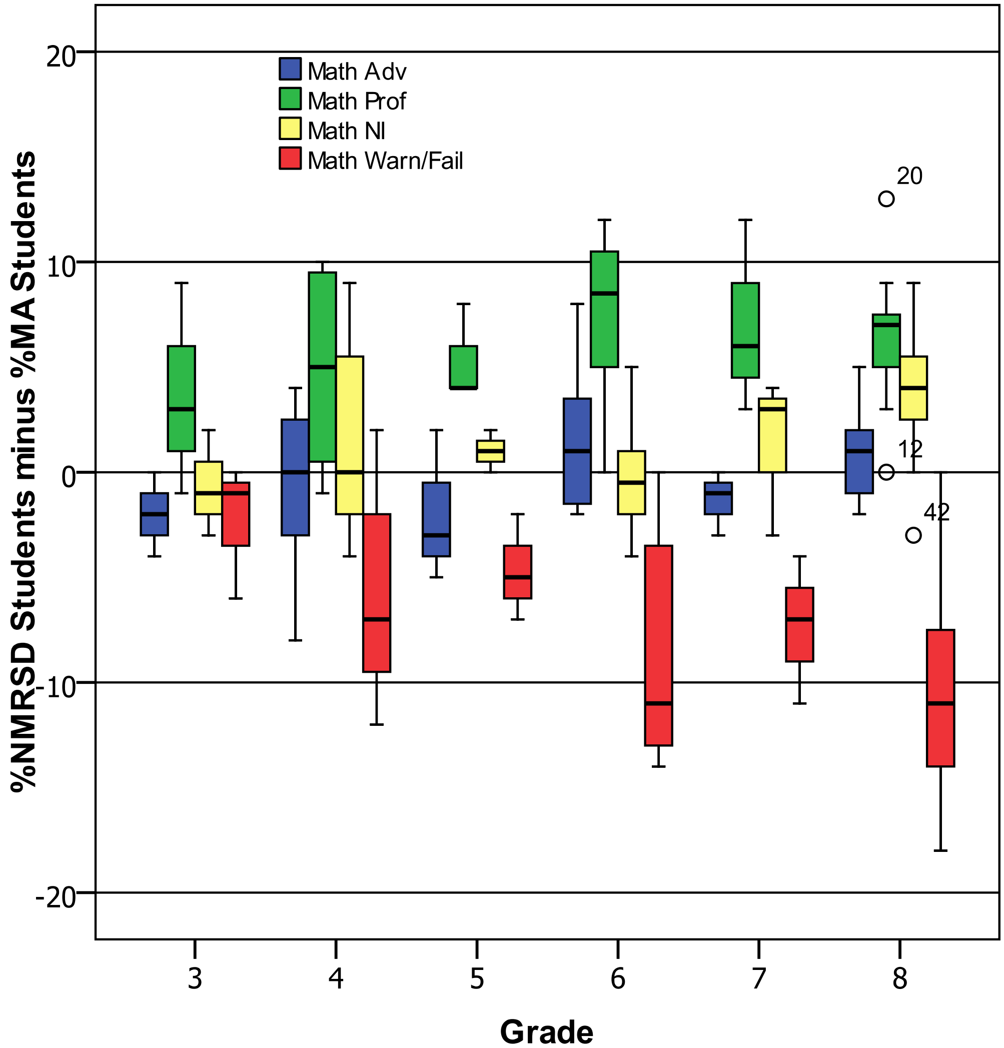


Figure 8. Percentage of NMRSD students minus Percentage of MA students by Mathematics Proficiency Levels: by Grade

A look at MCAS results in years preceding SM implementation is informative, and MCAS mathematics results for 1998 are shown in **Figure 9**. These are typical for the period and examples for 1999, 2000, 2004 and 2008 are shown in Appendix 1 – Univariate Analyses beginning on page 47.

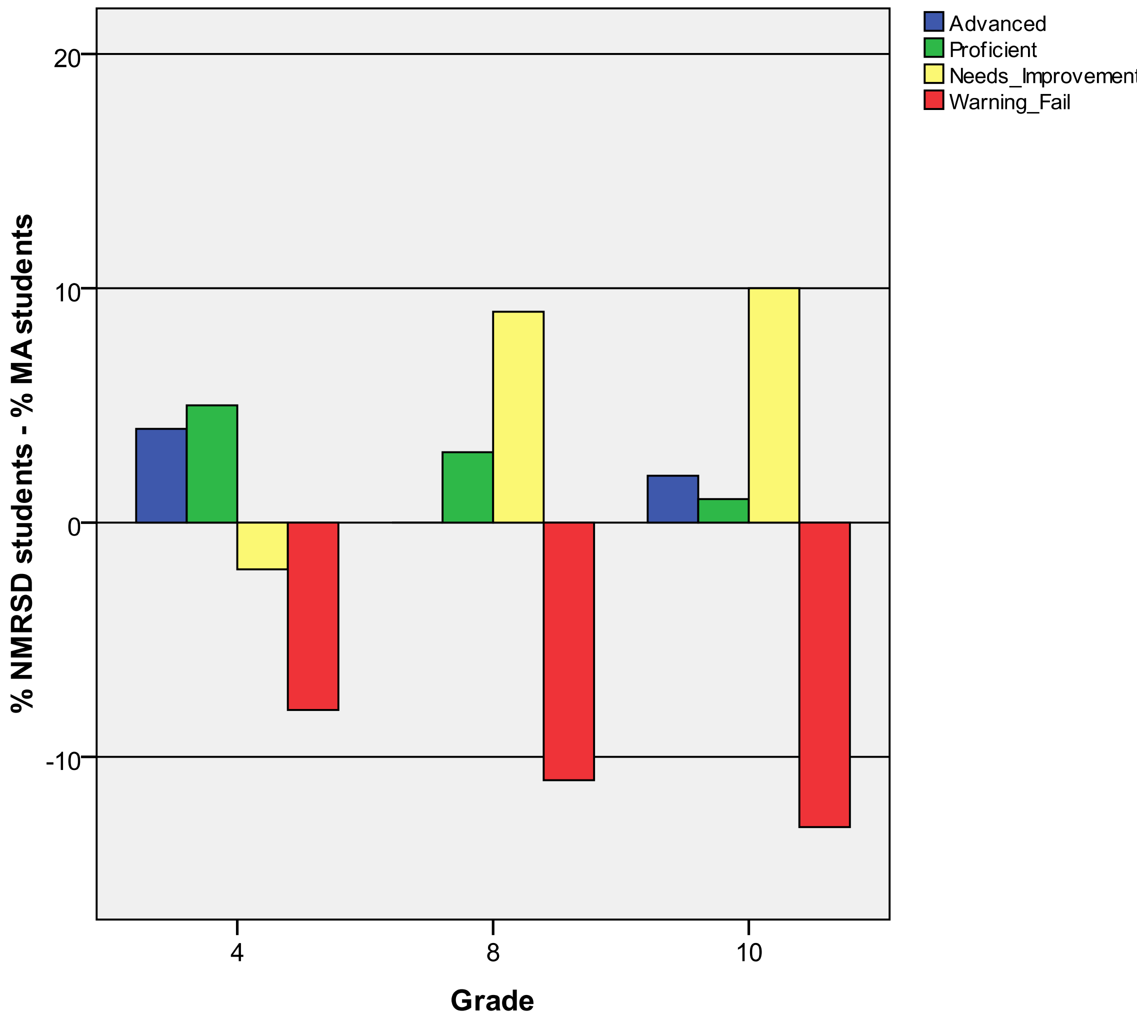


Figure 9. Differences, by Categorical Math Data, between NMRSD and MA in the 1998 MCAS.

This relation remains the norm for MCAS scores for the years 1998-2009. That is, both before and after the SM implementation, the proportion of NMRSD students with good proficiency levels is higher than that of their MA contemporaries.

Finally, Figure 14 on page 63 shows the relative percentage of students in the proficiency categories for NMRSD compared both to MA and to Quabbin School District, a district similar to NMRSD.

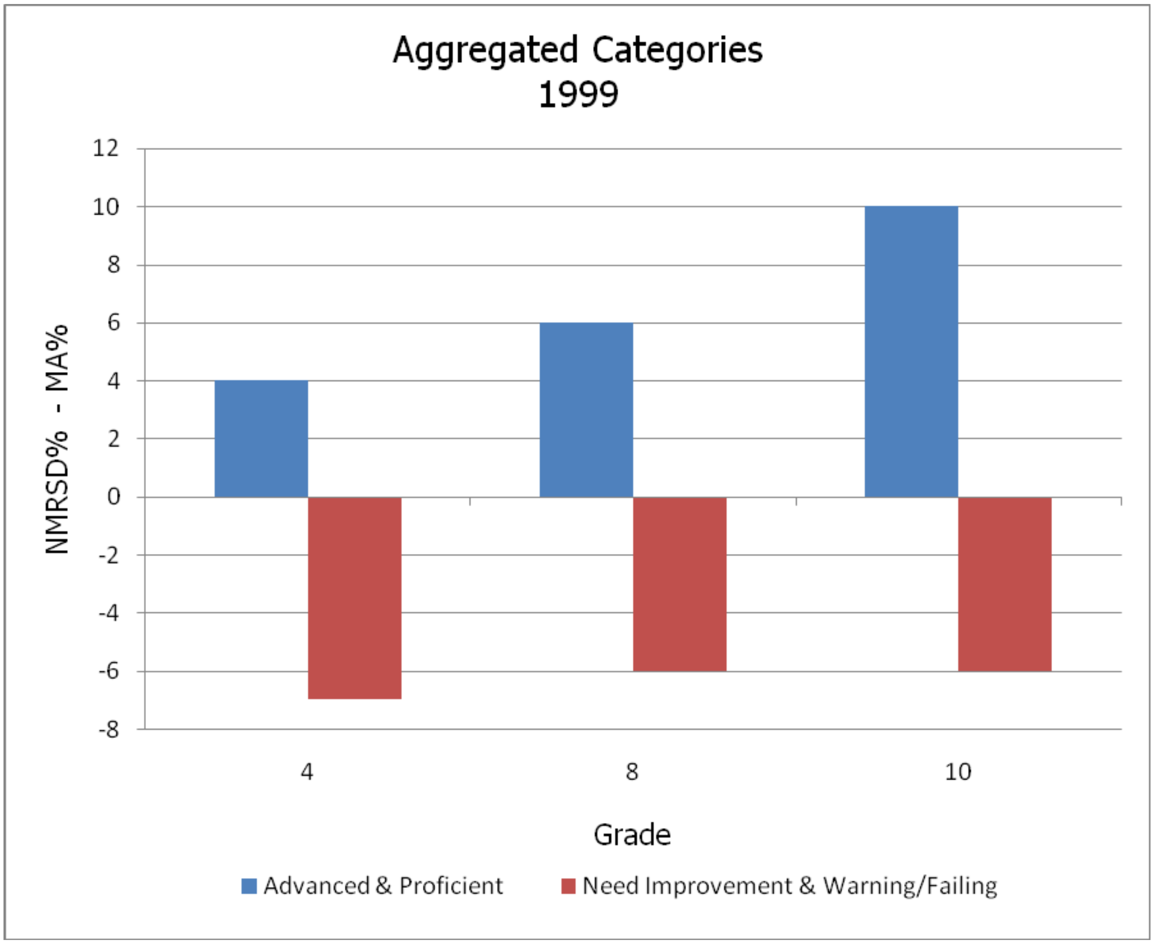


Figure 10. Aggregated Categories for MCAS Categorical Results, 1999.

Contingency table analyses for each grade-year for the SM vectors are presented in Table 25 in Appendix 3 – χ^2 on categories on MCAS, beginning on page 64. While most of these are significant for grade-year and while mean scores uniformly rise with increased CSM, chi-square, regression and all other relational measures between category and CSM were almost never significant; 7th grade in cohort V3 has, for example, x^2 = 12.97 (df 6, P=0.044).

A number of suggestive trends are apparent, since with the exception of 4th grade scores, MA and NMRSD scores do seem to follow the same general trends. NMRSD results are normally higher, see particularly Figure 12 and Figure 13 on pages 48 and 49, respectively.

Multivariate Methods

INTRODUCTION

We utilized regression with either of two models for both SM and CSM: one relating only SM or CSM to MCAS scores (Model 1) and another relating SM or CSM, Teacher, Free Lunch and Special Ed and MCAS scores (Model 2). SM analyses were binary with SM either 0 or 1 for the particular student, grade-year and MCAS score. In addition, we utilized generalized linear modeling for multivariate and repeated measures analyses. We analyzed all cases. Note that vectors V1 and V4 are the only vectors that have sufficient range of CSM, including baseline no-SM grade-years that are necessary to allow meaningful analyses. These analyses are intended to describe the relationship of CSM to MCAS results. Only V1 is fully presented in this report; results for V4 are similar, keeping in mind that only one year in V4 has 0 SM and that for only about 80 students.

Significant results suggest that means for the MCAS scores were not the same for each level of the tested variable, i.e.,

$$H_0: \ \mu_{teachers_a} = \mu_{teachers_b}$$

$$H_0: \ \mu_{CSM_0} = \mu_{CSM_1}$$

and so forth.

Regressions MCAS Scores by SM for all data

For all data for Model 2, Free Lunch was t=-8.563 with coefficient -2.781 and P=0.000, Special Ed was t=-35.902 with coefficient -10.268 and P=0.000, Teacher was t=3.870 with coefficient 0.008 and P=0.000 and SM was t=3.729 with coefficient 1.324 and P=0.000.

Over the period 2002 – 2009, there are 114 teachers in the 1-8 part of our data set, each of whom taught SM to as many as 429 students. To correct for teachers with only a few students, we also analyzed a subset of 44 teachers, each with more than 60 students. For this subset, Free Lunch was t=-8.290 with coefficient -3.102 and P=0.000; Special Ed was -28.383 with coefficient -9.729 and P=0.000; and Teacher was t=2.218 with coefficient 0.006 and P=0.027.

Data for teachers with >60 students by SM is presented at Appendix 2 – Individual Teacher Results, General Linear Model beginning on page 64. For these data and this model, SM had a positive effect on MCAS raw scores – students with no SM had mean scores 1.309 points below students with SM. Students in the Free

Lunch program scored 3.033 points below those not in the program and special education students scored 9.717 points below those not in special ed. Most teacher results were also negative. Reflecting the coding (no SM=0 and SM=1), most teacher coefficients are negative – i.e., MCAS raw scores for these teachers are higher than the No-SM mean MCAS score.

Regressions MCAS Scores by CSM for vector 1-4

For Model 1 for all vectors, the CSM t-score was 2.231 with coefficient 0.200 and P=0.026. The known CSM range for the period analyzed was 0-6, though the nominal CSM for 2002 – 2009 could have ranged from 0-8.

For V1, Model 2, the CSM t=3.729 with coefficient 1.324 and P=0.000; Special Ed t=-8.760 with coefficient -11.212 and P=0.000; Free Lunch t=-8.760 with coefficient -4.093 and P= 0.000 while Teacher t=3.870 with coefficient 0.08 and P=0.000.

For V2, CSM is without baseline, i.e., there are no non-SM scores.
For Model 2:
For 2006 Free Lunch t=3.827, coefficient 5.499 with P=0.000; Special Ed with t=-4.832, coefficient -7.7023 and P=0.000 then Teacher t=3.049, coefficient 1.262 and P=0.002.
For 2007, Free Lunch t=2.869 with coefficient 4.816 and P=0.004; Special Ed t=-4.937, coefficient -8.541 and P=0.000 and Teacher t=-3.452, coefficient -0.707 and P=0.001.
For 2008, Free Lunch t=3.679, coefficient 6.884 and P=0.000, Special Ed t=-5.341 with coefficient -10.793 and P=0.000 and, finally, Teacher t=0.361 with coefficient 0.075 and P=0.718.

For V3, all SM are "yes" so we have no baseline; there are no Free Lunch data for 2006 and there are no Teacher data for 2008, so the models are different each year; Special Ed and Teacher for 2006, Model 2 less CSM for 2007 and Special Ed and Free Lunch for 2008. The results are t=-5.446 with coefficient -9.023 and P=0.000 for Special Ed and t=4.029 with coefficient 1.230 and P=0.000 for Teacher.
For 2007 Special Ed is t=-3.904, coefficient =-5.374 and P=0.000; Free Lunch is t=3.479, coefficient 4.134 and P=0.001.
Note that here in 2007, Teacher t=-1.464, coefficient -0.149 and P=0.144, and is the only instance where Teacher was not a significant factor.
Finally, for 2008, Free Lunch t=4.438, coefficient 5.482 and P=0.000 and Special Ed t=-4.470 with coefficient -6.502 and P=0.000; there are no Teacher data for 2008.

V4 is, like V1, a vector for which we have better data, that is, data with few elisions.
For Model 2:
In 2005, CSM had t = -1.054 with coefficient -1.289 and P = 0.293; Free Lunch t= 2.087, coefficient 3.158 and P = 0.038; Special Ed t= -9.428, coefficient -14.130 and P = 0.000; and, finally, Teacher t = -0.639, coefficient -0.103 and P = 0.524.
In 2006, CSM was t = -3.133 with coefficient -4.129 and P = 0.002; Free Lunch t= 2.917, coefficient 5.20 and P = 0.004; Special Ed t = -7.403, coefficient -12.569 and Pe = 0.000; and Teacher t = 0.050, coefficient 0.007 and P = 0.961.
In 2007, CSM t = -1.059 with coefficient -1.495 and P = 0.291; Free Lunch t= 1.898, coefficient 2.919 and P = 0.059; Special Ed t = -8.699, coefficient= -12.954 and P = 0.000; and Teacher t = 2.559, coefficient 0.448 and P = 0.011.
In 2008, we have no Teacher data but CSM t = 0.791 with coefficient 0.858 and P = 0.430; Free Lunch t = 2.194, coefficient 3.393 and P = 0.029 and Special Ed t = -8.358, coefficient -12.524 and P = 0.000.
For 2009, CSM t = 1.493 with coefficient 1.933 and P = 0.137; Free Lunch t = 2.093, coefficient 3.882 and P = 0.038; Special Ed t = -7.368, coefficient -13.297 and P = 0.000; and Teacher t = -1.534, coefficient -0845 and P = 0.127.

General Linear Model for Multivariate statistics

All Cases

For all data, about 14,036 data points from 2002 to 2009, we used a binary representation of SM, i.e., either yes or no. In the analyses we compared MCAS raw scores for math and the five other sub-categories of mathematical understanding in the MCAS lexicon: Number Sense, Pattern Recognition, Geometry, Measurement and Data, Statistics and Probability. Results for this analysis are shown at Table 12 and Table 13, below. For these analyses the model has the MCAS subcategories above as dependent variables and the model effects are Intercept + SM + teacher + Free Lunch + Special Ed + Free Lunch * Special Ed. As may be seen, with the exception of the cross-product of Free Lunch and Special Ed.; all the effects are significant.

For all cases: only a relatively small proportion of the variability of the results can be associated with the model effects; R^2's range from about 0.10 to 0.20. While this is to be expected in such large, amorphous data sets, we believe that diachronic analyses may provide more substantive suggestions of relationships. See also Individual Grade-Years on page 43 for an example of the difficulties with this.

Table 12. Multivariate Tests, All Cases						
Effect	Statistic	Value	F (exact)	Hypothesis df	Error df	Sig.
Intercept	Hotelling -Lawley Trace	0.098	136.854	5	7013	0.000
SM	Hotelling -Lawley Trace	0.003	4.400	5	7013	0.001
teacher	Hotelling -Lawley Trace	0.043	60.645	5	7013	0.000
Free Lunch	Hotelling -Lawley Trace	0.013	18.178	5	7013	0.000
Special Ed	Hotelling -Lawley Trace	0.105	147.465	5	7013	0.000
Free Lunch * Special Ed	Hotelling -Lawley Trace	0.001	1.356	5	7013	0.238

Table 13. Multivariate Analysis of SM, Teacher, Free Lunch and Special Ed for All Data – Between Subjects Effects				
Source	Dependent Variable	df	F	Sig.
SM	MATHEMATICS Raw Score	1	12.225	0
	Number Sense Raw Score	1	5.543	0.019
	Patterns, Relations and Algebra Raw Score	1	9.935	0.002
	Geometry Raw Score	1	15.567	0
	Measurement Raw Score	1	1.103	0.294
	Data, Statistics and Probabilty Raw Score	1	5.452	0.02
teacher	MATHEMATICS Raw Score	1	10.234	0.001
	Number Sense Raw Score	1	69.083	0
	Patterns, Relations and Algebra Raw Score	1	39.977	0
	Geometry Raw Score	1	1.7	0.192
	Measurement Raw Score	1	6.667	0.01
	Data, Statistics and Probabilty Raw Score	1	99.508	0
Free Lunch	MATHEMATICS Raw Score	1	71.554	0
	Number Sense Raw Score	1	34.201	0
	Patterns, Relations and Algebra Raw Score	1	48.558	0
	Geometry Raw Score	1	41.224	0
	Measurement Raw Score	1	15.992	0
	Data, Statistics and Probabilty Raw Score	1	57.642	0
Special Ed	MATHEMATICS Raw Score	1	701.611	0
	Number Sense Raw Score	1	423.207	0
	Patterns, Relations and Algebra Raw Score	1	429.885	0
	Geometry Raw Score	1	256.796	0
	Measurement Raw Score	1	314.219	0

Table 13. Multivariate Analysis of SM, Teacher, Free Lunch and Special Ed for All Data – Between Subjects Effects				
Source	Dependent Variable	df	F	Sig.
	Data, Statistics and Probabilty Raw Score	1	423.743	0
Free Lunch * Special Ed	MATHEMATICS Raw Score	1	2.496	0.114
	Number Sense Raw Score	1	1.398	0.237
	Patterns, Relations and Algebra Raw Score	1	5.12	0.024
	Geometry Raw Score	1	0.914	0.339
	Measurement Raw Score	1	0.094	0.76
	Data, Statistics and Probabilty Raw Score	1	0.275	0.6

V1- 4th grade 2002 – 8th grade 2008

Multivariate Analyses

For V1 the model for multivariate analyses is 04_2002_Math_Raw_Score, 06_2004_Math_Raw_Score, 08_2006_Math_Raw_Score with 06_2004_CSM, and 08_2006_CSM (where for example 04_2002 refers to 4th graders tested in 2002). Results are shown in Table 14 and Table 15, below. Table 14 gives the Wilk's Lambda for the effect (CSM) for V1 and, while all but 4th grade in 2002 are "significant", these analyses, again, do not account for most of the variability in the data sets – the partial η^2 translate to R^2's of about 0.2, 0.005, 0.06 and 0.095 for the three levels of CSM.

Table 14. Multivariate tests for V1							
Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Wilks' Lambda	0.815	22.701	3.000	301	0.000	0.185
CSM, 4th grade	Wilks' Lambda	0.994	0.619	3.000	301	0.603	0.006
CSM, 6th grade	Wilks' Lambda	0.937	6.698	3.000	301	0.000	0.063
CSM, 8th grade	Wilks' Lambda	0.905	10.575	3.000	301	0.000	0.095

Table 15. Between-subjects Effects, V1						
Source	Dependent Variable	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	4th grade Math Raw Score	10	350.112	9.205	.000	0.138
	6th grade Math Raw Score	10	317.584	12.324	.000	0.129
	8th grade Math Raw Score	10	4.423	10.948	.000	0.130
CSM, 4th grade	4th grade Math Raw Score	1	7.566	0.102	0.750	0.000
CSM, 6th grade	6th grade Math Raw Score	2	88.264	1.221	0.270	0.004
CSM, 8th grade	8th grade Math Raw Score	4	615.979	6.184	0.000	0.077

Table 15 shows the effects of CSM on the Math Raw Scores for V1. Again, while "significant" the 8th grade, results still suffer from the high residual sum of squares problem where little of the variability in the data are accounted for by the model.

Repeated measures

For these analyses, we use Math Raw scores for 4th, 6th and 8th grades as the measure and CSM for the between-subjects factors, or 4th grade (2002), 6th grade (2004) and 8th grade (2006) Math raw scores by 4th grade, 6th grade and 8th grade CSM. The results for these are similar to the preceding, with very large residual sums of squares limiting our confidence in the significance of the results.

Table 16. Repeated-measures tests V1							
Effect	Test	Value	F (exact)	Hypothesis df	Error df	Sig.	Partial Eta Squared
factor1	Wilks' Lambda	.966	5.251	2	295	.006	.034
factor1 * @04_2002_CSM	Wilks' Lambda	.992	1.249	2	295	.288	.008
factor1 * @06_2004_CSM	Wilks' Lambda	.989	1.576	2	295	.208	.011
factor1 * @08_2006_CSM	Wilks' Lambda	.939	2.371	8	590	.016	.031
factor1 * @04_2002_CSM * @06_2004_CSM	Wilks' Lambda	1.000	.		296	.	.

Table 16. Repeated-measures tests V1							
Effect	Test	Value	F (exact)	Hypothesis df	Error df	Sig.	Partial Eta Squared
factor1 * @04_2002_CSM * @08_2006_CSM	Wilks' Lambda	1.000	.006	2	295	.994	.000
factor1 * @06_2004_CSM * @08_2006_CSM	Wilks' Lambda	1.000	.		296	.	.
factor1 * @04_2002_CSM * @06_2004_CSM * @08_2006_CSM	Wilks' Lambda	1.000	.		296	.	.

Table 17 shows the Levene's test for the equality of error variances; in this case, since none are significant, we cannot say that variances are different from one other, and we may place some reliance on differences and allow ourselves to look at the means for each of the test years by CSM.

Table 17. Levene's Test of Equality of Error Variances				
	F	df1	df2	Sig.
4^{th} grade Math Raw Score	1.719	10	296	.076
6^{th} grade Math Raw Score	.886	10	296	.546
8^{th} grade Math Raw Score	1.167	10	296	.313

In Table 18 we can see the cumulative effect of SM on mean MCAS Raw scores; results for sub-categories are similar. To illustrate the table, we'll look at a couple of examples.

1) For 4^{th} grade MCAS Raw scores, students with no SM had a mean score of 25.5 in 4^{th} grade, 21.8 in 6^{th} grade and 19 in 8^{th} grade.
2) Students with no SM in 4^{th}, 5^{th} and 6^{th} grade but SM in 7^{th} and 8^{th} grades had a mean score of 34.4 in 4^{th} grade, 39 in 6^{th} grade and 37 in 8^{th} grade. In addition to the effect of CSM, this shows how the propinquity in time of SM affects results. This is reinforced by the results of students with 1 year of SM by 6^{th} grade and 3 years of SM by 8^{th} grade (i.e., SM in 7^{th} and 8^{th}

grades) with MCAS mean score of 37.1 in 4th grade, 40.5 in 6th grade and 39.9 in 8th grade.

3) A final example – the students with SM in each year, i.e., 4th through 8th grades, had means of 44.6, 43.9 and 44.8.

These results clearly show the effect of CSM over the period covered in the vector V1.

Table 18. Means by CSM History for V1

Grade-Score	CSM 04-2002	CSM 06-2004	CSM 08-2006	Mean	Std. Dev.	N
4th Grade MCAS Math RAW	0	0	0	25.5	7.234	4
		0	2	34.4	8.484	142
		1	1	11.5	7.778	2
		1	3	37.1	8.949	54
		2	3	36.0	.	1
		2	4	35.7	8.351	78
	1	1	1	19.0	.	1
		1	3	30.3	17.898	3
		2	2	28.0	.	1
		2	4	32.1	12.335	7
		3	5	44.6	5.611	14
6th Grade MCAS Math RAW	0	0	0	21.8	8.732	4
		0	2	39.0	8.173	142
		1	1	19.0	7.071	2
		1	3	40.5	9.691	54
		2	3	50.0	.	1
		2	4	41.0	8.062	78
	1	1	1	32.0	.	1
		1	3	39.3	8.737	3
		2	2	20.0	.	1
		2	4	39.1	9.839	7
		3	5	43.9	8.678	14
8th Grade MCAS Math RAW	0	0	0	19.0	6.733	4
		0	2	37.0	9.886	142
		1	1	12.5	2.121	2
		1	3	39.9	10.924	54
		2	3	49.0	.	1
		2	4	38.2	9.778	78
	1	1	1	26.0	.	1
		1	3	40.0	10.817	3
		2	2	19.0	.	1
		2	4	35.1	10.511	7
		3	5	44.8	8.684	14

Individual Grade-Years

The difficulty of post-hoc analysis of historical data is illustrated by Table 19 which shows results for 6th grade in 2005. The Model is CSM+MCAS Raw score +

each of the sub-categories of MCAS Math. While the model as a whole accounts for much of the variability in the set – illustrated by the R^2 – CSM and the other factors are much less reliable as predictors. For example, the sum of squares (Type III) for CSM is 2,614 and the error ss is 26,495 for this model and grade-year.

Table 19. GLM Between-Subjects Effects for one grade-year

Source	Dependent Variable	df	Mean Square	F	Sig.	R^2
Model	6th grade Math Raw Score	4	128334.8	1682.684	0	0.950
	6th grade Number Sense	4	14842.34	1752.012	0	0.952
	6th grade Patterns & Relationships	4	9582.826	1414.576	0	0.942
	6th grade Geometry	4	2281.472	648.321	0	0.881
	6th grade Measurement	4	1562.6	479.589	0	0.845
	6th grade Data, Statistics & Probability	4	2655.045	935.502	0	0.914
@06_2005_CSM	6th grade Math Raw Score	1	2613.975	34.274	0	0.09
	6th grade Number Sense	1	333.438	39.359	0	0.10
	6th grade Patterns & Relationships	1	179.796	26.541	0	0.07
	6th grade Geometry	1	35.331	10.04	0.002	0.03
	6th grade Measurement	1	33.933	10.414	0.001	0.03
	6th grade Data, Statistics & Probability	1	59.118	20.83	0	0.06
@06_2005_teacher	6th grade Math Raw Score	1	709.01	9.296	0.002	0.03
	6th grade Number Sense	1	62.337	7.358	0.007	0.02
	6th grade Patterns & Relationships	1	77.896	11.499	0.001	0.03
	6th grade Geometry	1	82.915	23.562	0	0.06
	6th grade Measurement	1	0.329	0.101	0.751	0.00
	6th grade Data, Statistics & Probability	1	1.888	0.665	0.415	0.00
Free Lunch	6th grade Math Raw Score	1	1116.93	14.645	0	0.04
	6th grade Number Sense	1	68.205	8.051	0.005	0.02
	6th grade Patterns & Relationships	1	33.81	4.991	0.026	0.01
	6th grade Geometry	1	43.913	12.479	0	0.03
	6th grade Measurement	1	43.718	13.418	0	0.04
	6th grade Data, Statistics & Probability	1	37.315	13.148	0	0.04
Special Ed	6th grade Math Raw Score	1	1780.888	23.35	0	0.06
	6th grade Number Sense	1	127.08	15.001	0	0.04
	6th grade Patterns & Relationships	1	53.58	7.909	0.005	0.02

Table 19. GLM Between-Subjects Effects for one grade-year						
Source	Dependent Variable	df	Mean Square	F	Sig.	R^2
	6th grade Geometry	1	101.682	28.895	0	0.08
	6th grade Measurement	1	44.247	13.58	0	0.04
	6th grade Data, Statistics & Probability	1	47.226	16.64	0	0.05

References

1. *TIMSS Results Place Massachusetts Amongh World Leaders in Math and Science*. 2008 09Dec09 [cited 2009 12 Sep 2009]; Available from: http://www.doe.mass.edu/news.asp?id=4457.
2. Ginsburg, A., et al., *What the United States Can Learn From Singapore's World-Class Mathematics System (and what Singapore can learn from the United States): An Exploratory Study*. 2005, American Institutes for Research®, 1000 Thomas Jefferson Street, NW, Washington, DC 20007-3835: Washington, DC
3. *2007 MCAS Technical Report*, in *Massachussets Comprehensive Assessment System*. 2007.
4. *2006 MCAS Technical Report*, in *Massachussets Comprehensive Assessment System*. 2006.
5. *2005 MCAS Technical Report*, in *Massachussets Comprehensive Assessment System*. 2006.
6. *2004 MCAS Technical Report*, in *Massachussets Comprehensive Assessment System*. 2005.
7. *2003 MCAS Technical Report*, in *Massachussets Comprehensive Assessment System*. 2003.
8. *2002 MCAS Technical Report*, in *Massachussets Comprehensive Assessment System*. 2003.
9. *2001 MCAS Technical Report*, in *Massachussets Comprehensive Assessment System*. 2002.
10. *2000 MCAS Technical Report*, in *Massachussets Comprehensive Assessment System*. 2002.
11. *1999 MCAS Technical Report*, in *Massachussets Comprehensive Assessment System*. 1999.
12. *1998 MCAS Technical Report Summary*, in *Massachussets Comprehensive Assessment System*. 1999.
13. *2008 MCAS Technical Report*, in *Massachussets Comprehensive Assessment System*. 2008.
14. Waight, M.M., *The Implementation of Singapore Mathematics in a Regional School District in Massachusetts:2000-2006*, in *Remarks to a National Mathematics Advisory Panel*. 2006: Cambridge, Massachusetts.

Appendix 1 – Univariate Analyses

Continuous Data

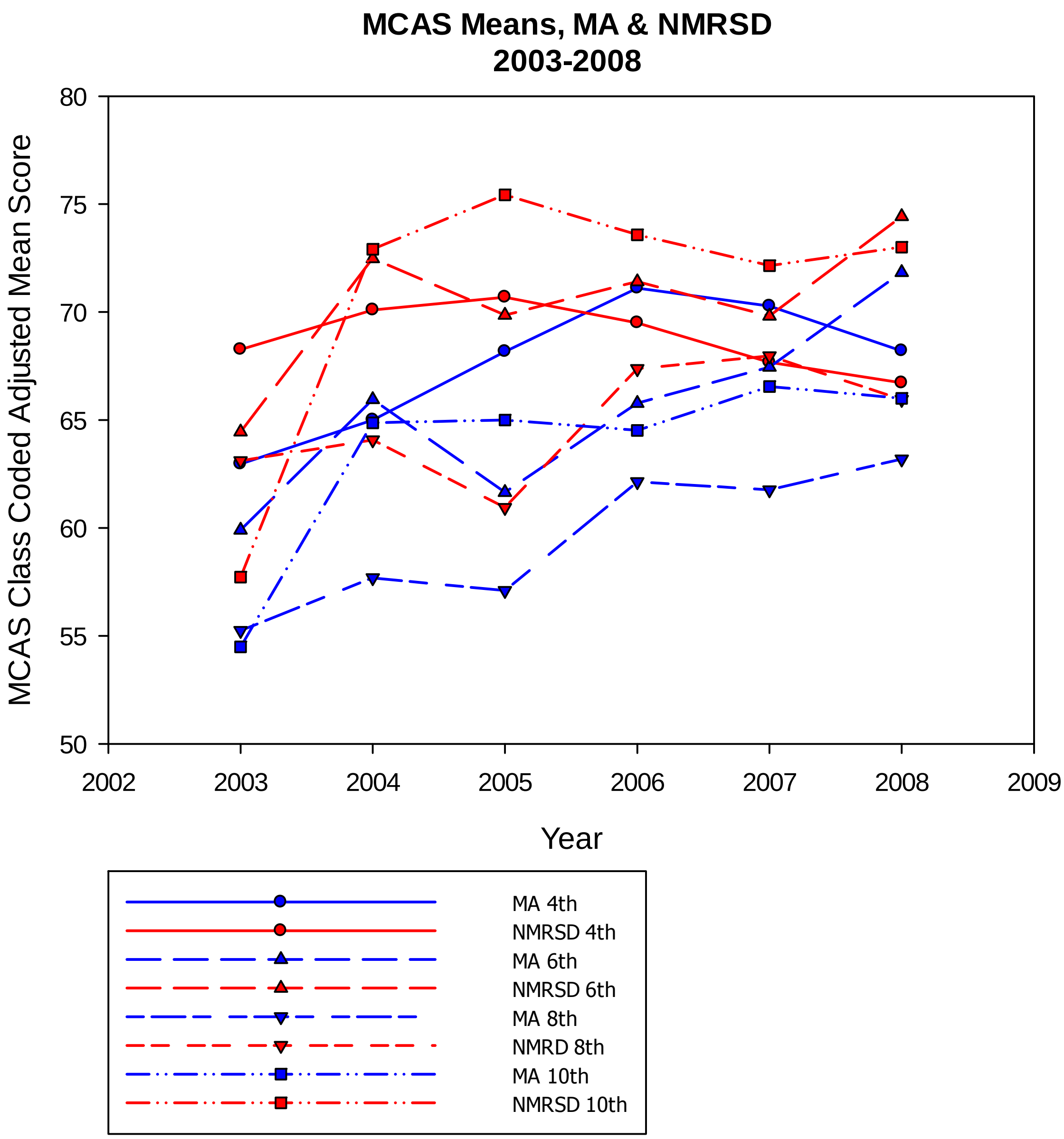


Figure 11. MCAS Mean Scores, all grades 2003-2008

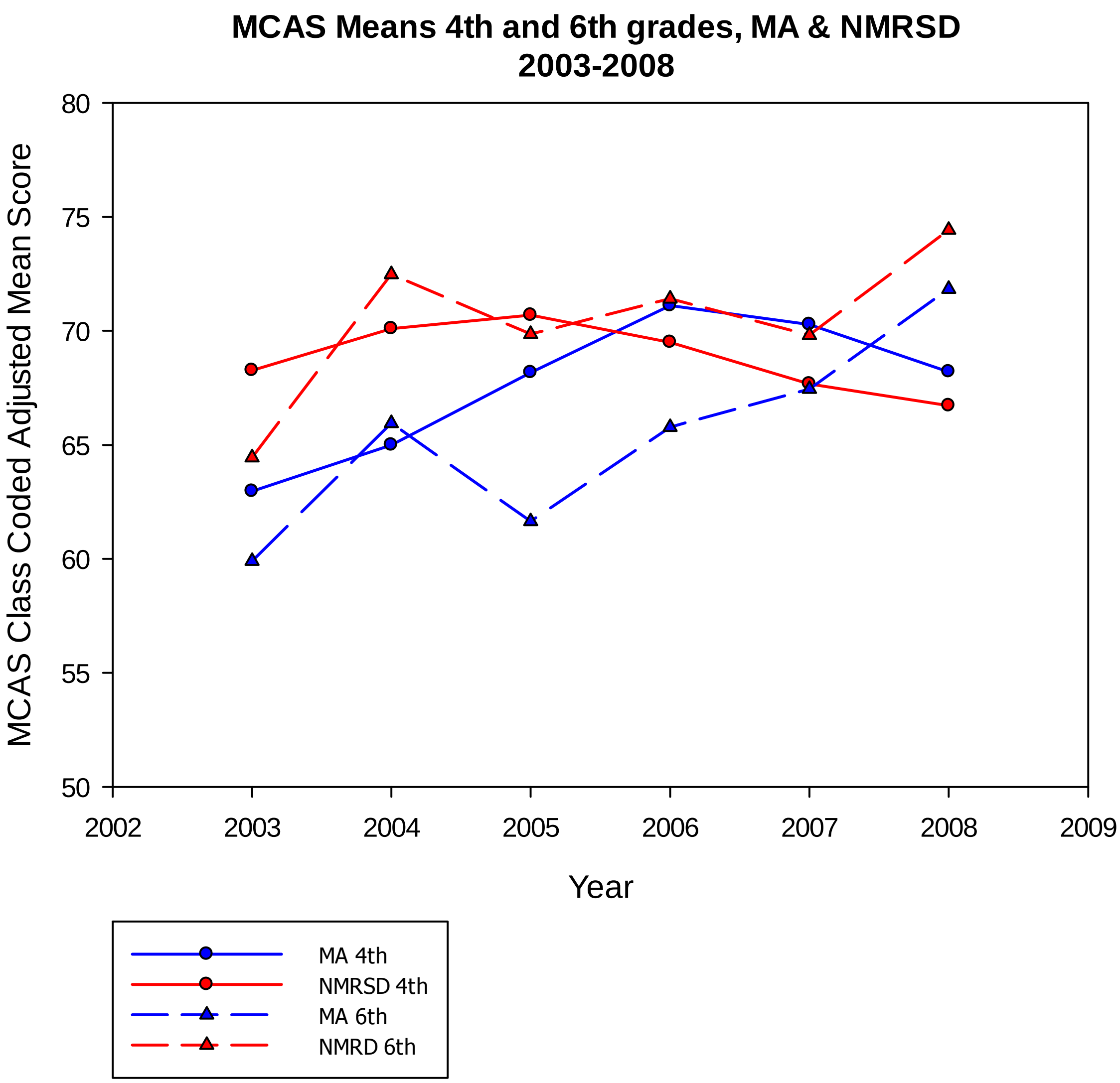


Figure 12. MCAS Means for 4th and 6th grades

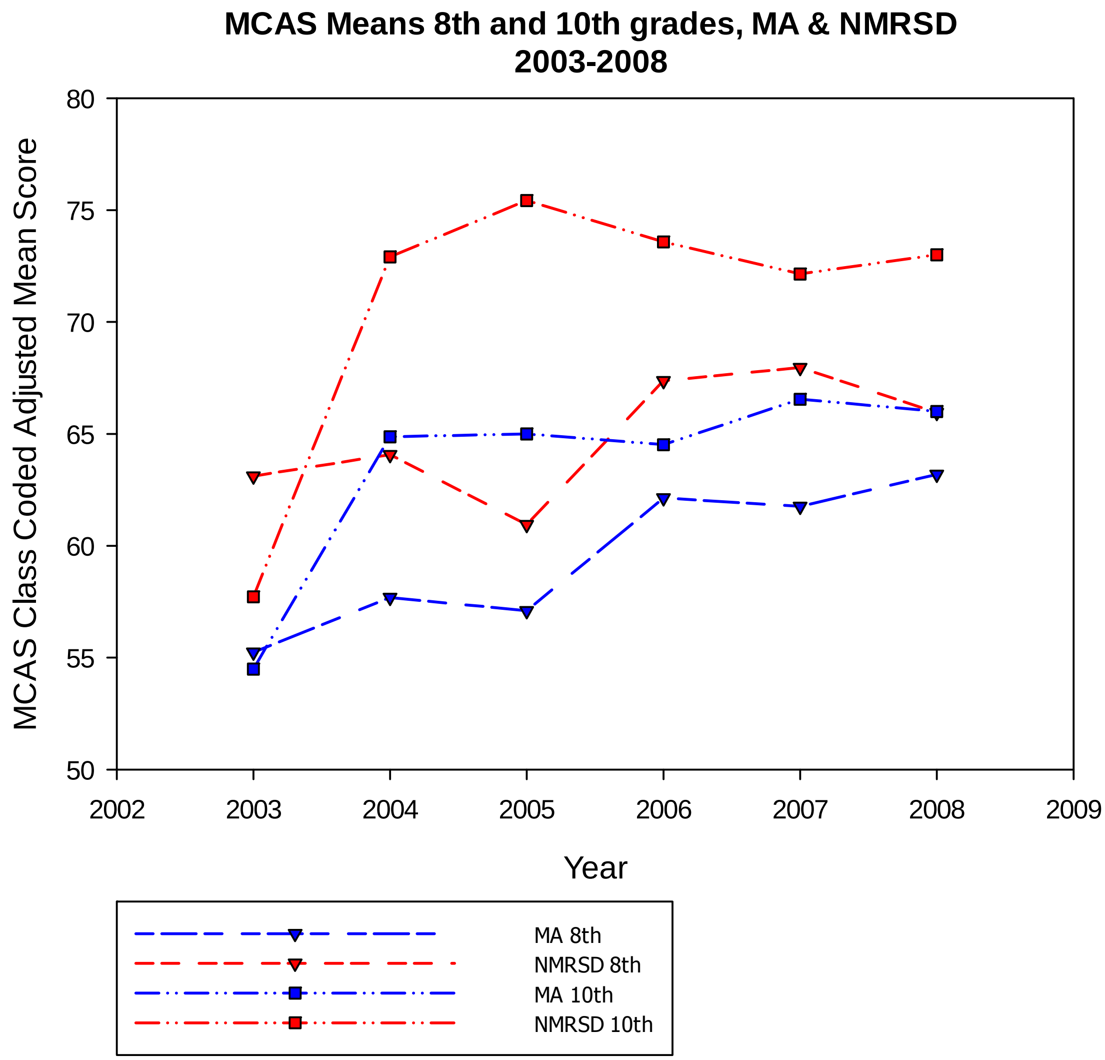


Figure 13. MCAS Mean Scores for 8th and 10th grades

Correlations on Continuous Data

The following tables and figures relate to consideration of continuous, or mean data for the years under study.

Table 20. Cohort Correlations between MCAS raw scores and CSM				
V1				
CSM=0-5	Grade 4	Grade 6	Grade 7	Grade 8
Spearman's rho	0.079	0.128	0.093	0.116
sig. (2-tailed)	0.156	0.021	0.095	0.037
N	325	327	325	323
V4				
CSM=0-4	Grade 4	Grade 6	Grade 8	
Spearman's rho	0.109	0.152	0.085	
sig. (2-tailed)	0.055	0.008	0.139	
N	307	307	307	

Table 21. Correlations between CSM and Social parameters				
V1				
Free Lunch				
Spearman's rho	-0.124	-0.126	-0.114	-0.146
sig. (2-tailed)	0.026	0.023	0.039	0.009
N	325	327	325	323
Gender				
Spearman's rho	-0.032	0.042	0.044	-0.004
sig. (2-tailed)	0.563	0.447	0.430	0.945
N	325	327	325	323
Special Education				
Spearman's rho	-0.342	-0.281	-0.205	-0.258
sig. (2-tailed)	0.000	0.000	0.000	0.000

Table 21. Correlations between CSM and Social parameters				
N	325	327	325	323
V4				
Free Lunch				
Spearman's rho	-0.054	-0.116	-0.110	
sig. (2-tailed)	0.346	0.042	0.054	
N	307	307	307	
Gender				
Spearman's rho	0.182	0.098	0.128	
sig. (2-tailed)	0.001	0.087	0.025	
N	307	307	307	
Special Education				
Spearman's rho	-0.327	-0.458	-0.406	
sig. (2-tailed)	.000	.000	.000	
N	307	307	307	

Table 22. Raw MCAS scores by CSM for Cohorts One-way ANOVA for Cumulative Singapore Math Experience for Raw MCAS scores					
Grade	CSM	N	Mean Raw	F, Raw Scores	Sig, Raw Scores
V1					
4	0	184	37.19	2.418	0.121
	1	141	38.72		
6	1	4	24.25	Excluded from ANOVA	
	2	180	37.91	4.996	0.026
	3	143	40.37		
7	1	2	30.50	Excluded	

Table 22. Raw MCAS scores by CSM for Cohorts One-way ANOVA for Cumulative Singapore Math Experience for Raw MCAS scores					
Grade	CSM	N	Mean Raw	F, Raw Scores	Sig, Raw Scores
	3	180	36.84	2.816	0.094
	4	143	38.69		
8	2	2	23.00	Excluded	
	4	178	35.71	3.647	0.052
	5	143	37.99		
V4					
4	0	281	35.02	2.558	0.111
	1	26	38.00		
6	0	146	38.58	2.293	0.078
	1	60	39.60		
	2	89	40.64		
	3	15	44.07		
7	0	3	19.67	Excluded	
	1	136	37.03	1.437	0.241
	2	52	39.96		
	3	77	38.51		
	4	15	44.33		
8	0	4	19.00	Excluded	
	1	145	38.91	0.646	0.586
	2	59	38.56		
	3	87	37.54		
	4	15	43.00		

Table 23. Effect of Free Lunch on MCAS scores by CSM One-way ANOVA for Cumulative Singapore Math Experience for Free Lunch						
Grade	CSM	N	Mean Raw Score	Free Lunch Yes/No	F, Free Lunch	Sig, Free Lunch
V1						
4	0	166	37.94	No	0.631	0.428
	1	129	38.74	No		
	0	18	30.28	Yes	5.101	0.032
	1	22	38.58	Yes		
6	2	162	38.69	No	2.716	0.1
	3	131	40.53	No		
	2	18	30.89	Yes	3.235	0.083
	3	12	38.67	Yes		
7	1	2	30.50	No	Excluded	
	3	162	37.67	No	2.816	0.094
	4	131	38.08	No		
	3	18	29.39	Yes	3.985	0.056
	4	12	38.50	Yes		
8	2	2	23.00	No	Excluded	
	4	160	36.63	No	1.614	0.205
	5	131	38.13	No		
	4	18	27.56	Yes	4.057	0.054
	5	12	36.42	Yes		
V4						
4	0	259	35.16	No	4.292	0.039
	1	24	39.08	No		
	0	22	33.32	Yes	1.038	0.319
	1	2	25.00	Yes		
6	0	137	39.01	No	1.972	0.118
	1	52	40.38	No		
	2	79	40.76	No		
	3	15	44.07	No		
	0	9	31.89	Yes	0.797	0.464
	1	8	34.50	Yes		
	2	7	39.29	Yes		
8	0	76	36.51	No	7.328	0.007
	1	99	38.65	No		
	2	35	39.89	No		
	3	67	38.03	No		

Table 23. Effect of Free Lunch on MCAS scores by CSM One-way ANOVA for Cumulative Singapore Math Experience for Free Lunch						
Grade	CSM	N	Mean Raw Score	Free Lunch Yes/No	F, Free Lunch	Sig, Free Lunch
	4	6	43.00	No		
	0	6	23.50	Yes	3.183	0.046
	1	9	41.78	Yes		
	2	4	27.00	Yes		
	3	5	31.00	Yes		

Table 24. Effect of Special Ed on MCAS scores by CSM One-way ANOVA for Cumulative Singapore Math Experience for Special Education						
Grade	CSM	N	Mean Raw Score	Sp Ed Yes/No	F, Sp Ed	Sig, Sp Ed
V1						
4	0	160	38.57	No	1.266	0.261
	1	127	39.67	No		
	0	24	28.00	Yes	0.636	0.430
	1	14	31.14	Yes		
6	2	160	39.18	No	3.406	0.066
	3	127	41.18	No		
	1	4	24.25	Yes	Excluded	
	2	20	27.75	Yes	2.754	0.106
	3	16	33.94	Yes		
7	3	160	37.64	No	2.948	0.087
	4	127	39.54	No		
	1	2	30.50	Yes	Excluded	
	3	20	30.45	Yes	0.147	0.704
	4	16	31.94	Yes		
8	4	158	36.98	No	2.023	0.156
	5	127	38.65	No		
	2	2	23.00	Yes	Ex-	

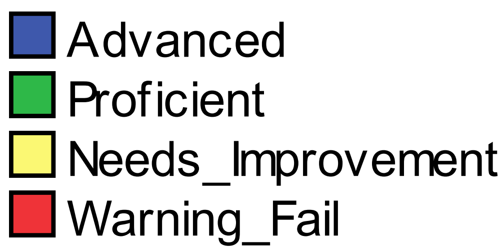
Table 24. Effect of Special Ed on MCAS scores by CSM One-way ANOVA for Cumulative Singapore Math Experience for Special Education						
Grade	CSM	N	Mean Raw Score	Sp Ed Yes/No	F, Sp Ed	Sig, Sp Ed
					cluded	
	4	20	25.65	Yes	3.629	0.065
	5	16	32.69	Yes		
V4						
4	0	254	35.90	No	12.943	0.000
	1	20	42.85	No		
	0	27	26.70	Yes	1.557	0.221
	1	6	21.83	Yes		
6	0	131	40.03	No	2.940	0.034
	1	50	42.66	No		
	2	78	42.19	No		
	3	15	44.07	No		
	0	15	25.87	Yes	0.126	0.882
	1	10	24.30	Yes		
	2	8	25.50	Yes		
8	0	74	36.89	No	2.417	0.049
	1	96	40.58	No		
	2	33	41.48	No		
	3	65	39.26	No		
	4	6	43.00	No		
8	0	8	23.25	Yes	0.286	0.835
	1	12	25.50	Yes		
	2	6	22.50	Yes		
	3	7	21.57	Yes		

Categorical Data

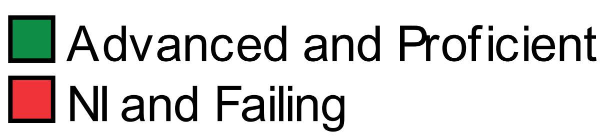
The following figures and tables relate to the comparison of categorical data, largely the "Proficiency Categories" of the MCAS.

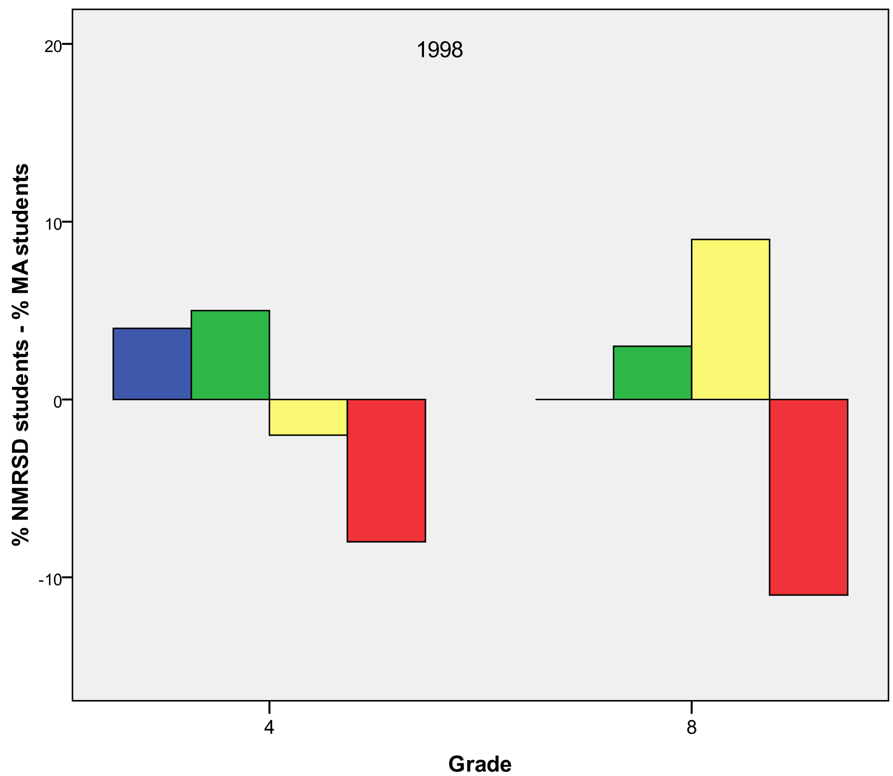
Graphs of Categories: NMRSD-MA percents of students
The following graphs are the percentage of students in each category from NMRSD minus the percentage of students in the corresponding category from MA state results. In other words, if 20% of NMRSD students were in the Advanced category in grade 5 for 2007 while MA had 14% of their students in the Advanced category, the difference for that category for that year was 6%.

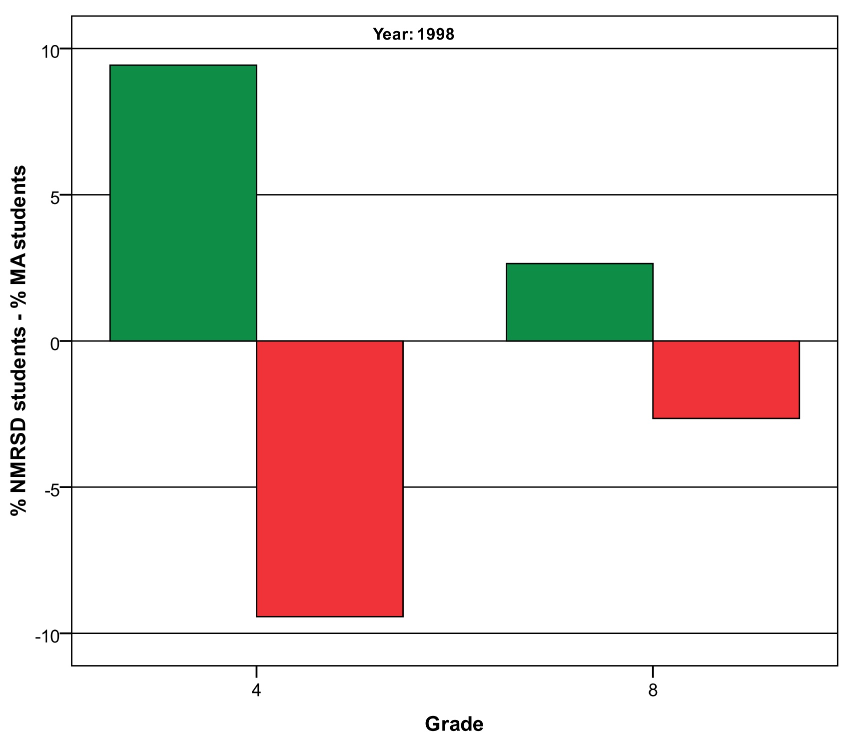
The following are the bar colors for the proficiency categories used in MCAS. These are used in the graphs in this section.

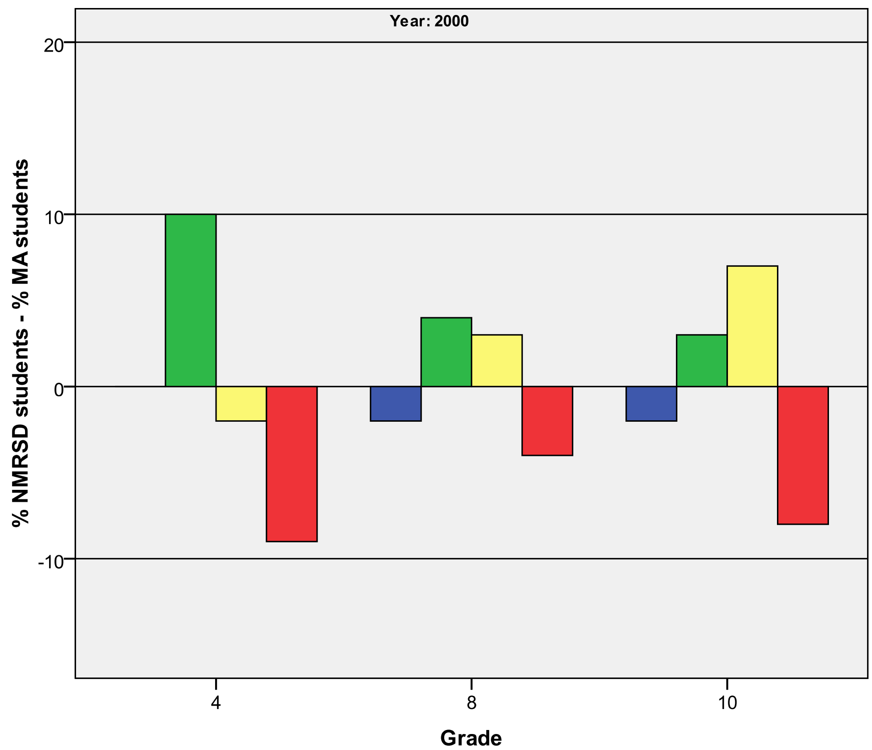


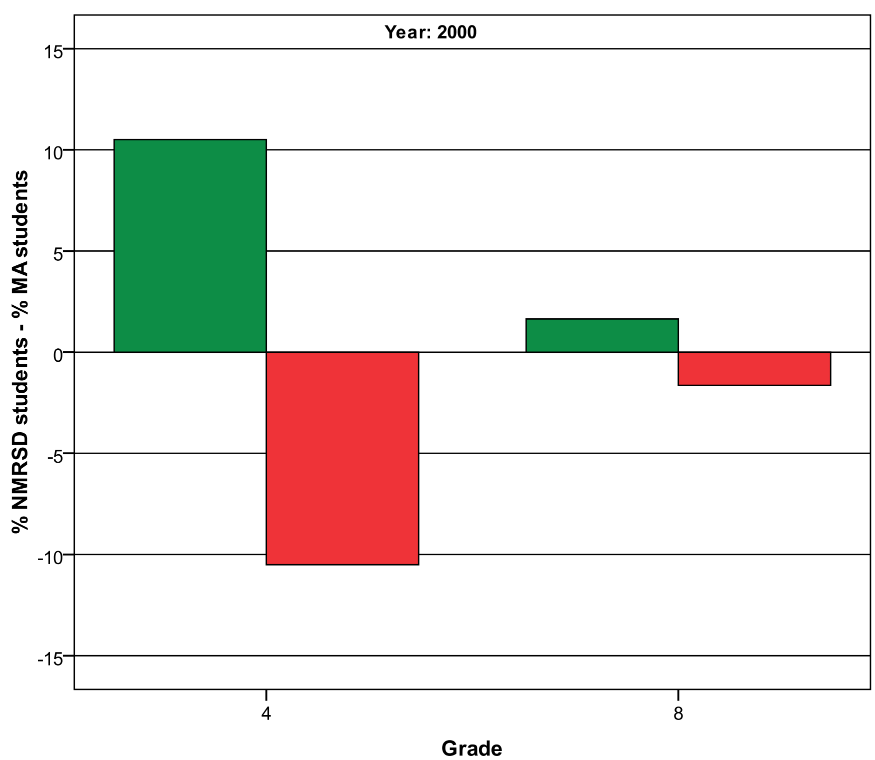
For the grouped categories:

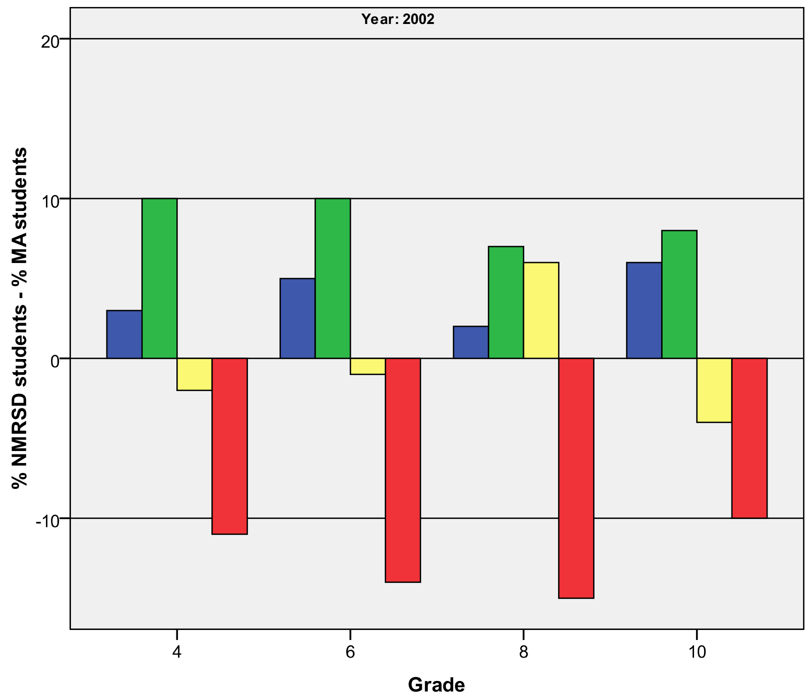


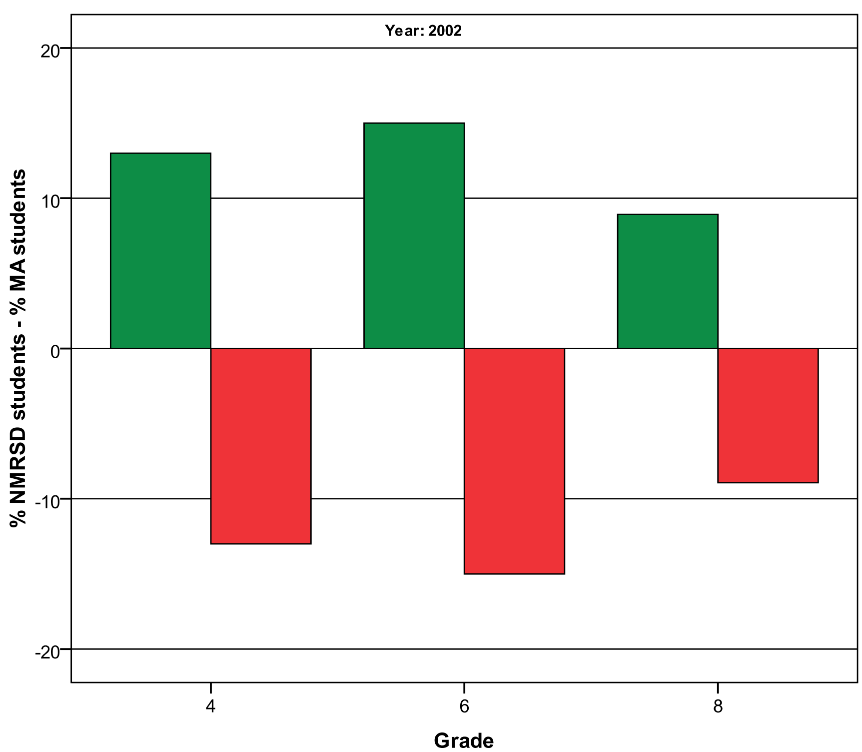

1998
20
10
0
-10
% NMRSD students - % MA students
4
8
Grade

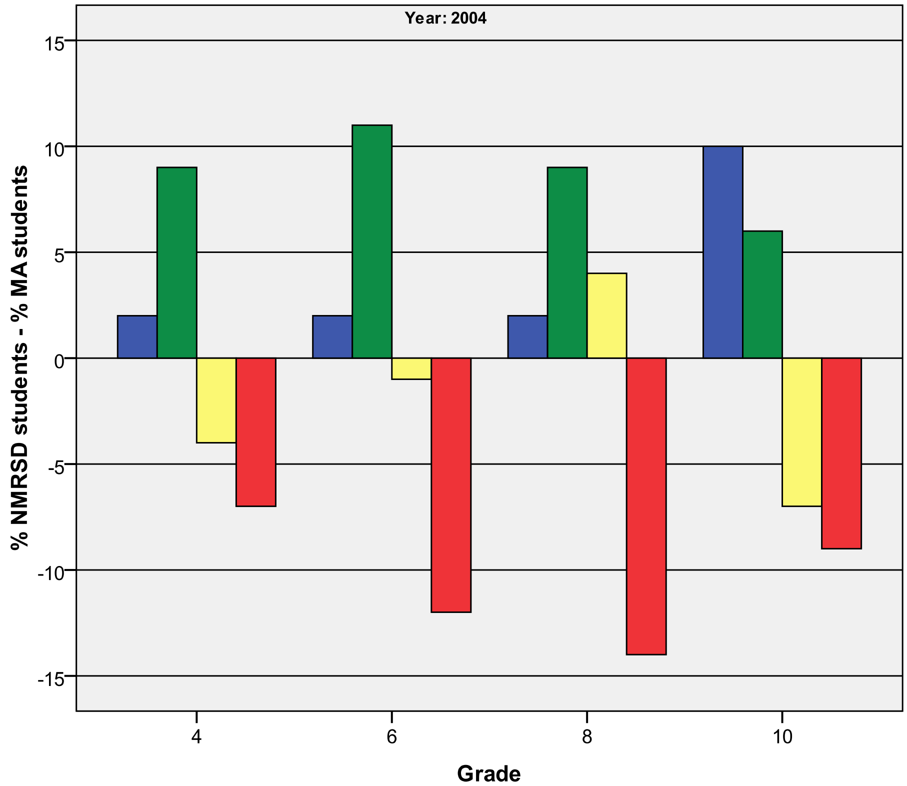

Year: 1998
10
5
0
-5
-10
% NMRSD students - % MA students
4
8
Grade

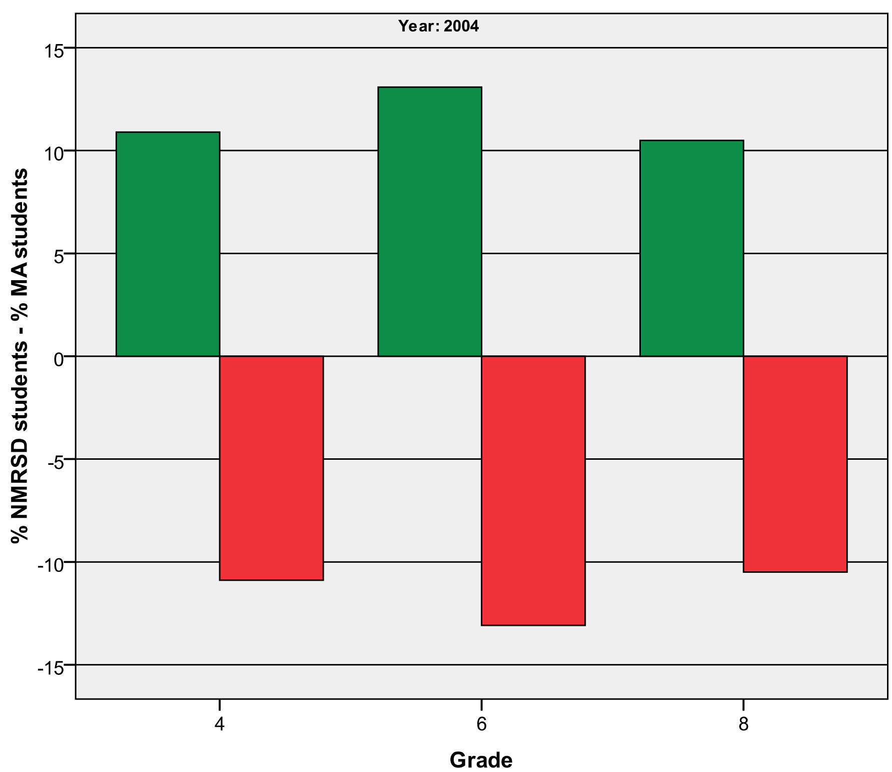

Year: 2000
20
10
0
-10
% NMRSD students - % MA students
4
8
10
Grade

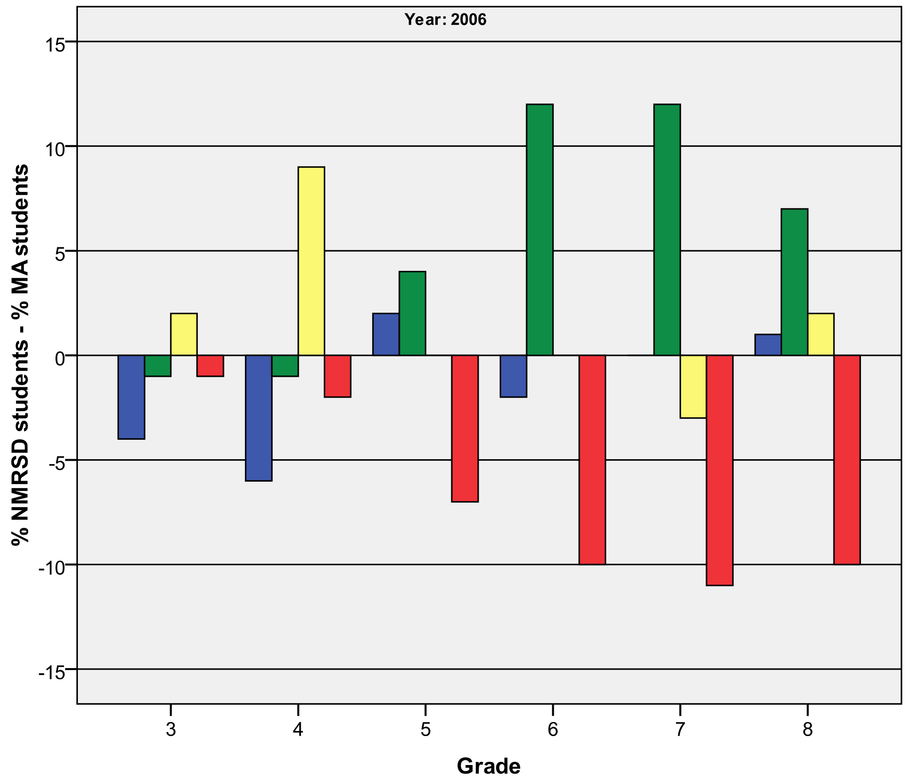

Year: 2000
15
10
5
0
-5
-10
-15
% NMRSD students - % MA students
4
8
Grade

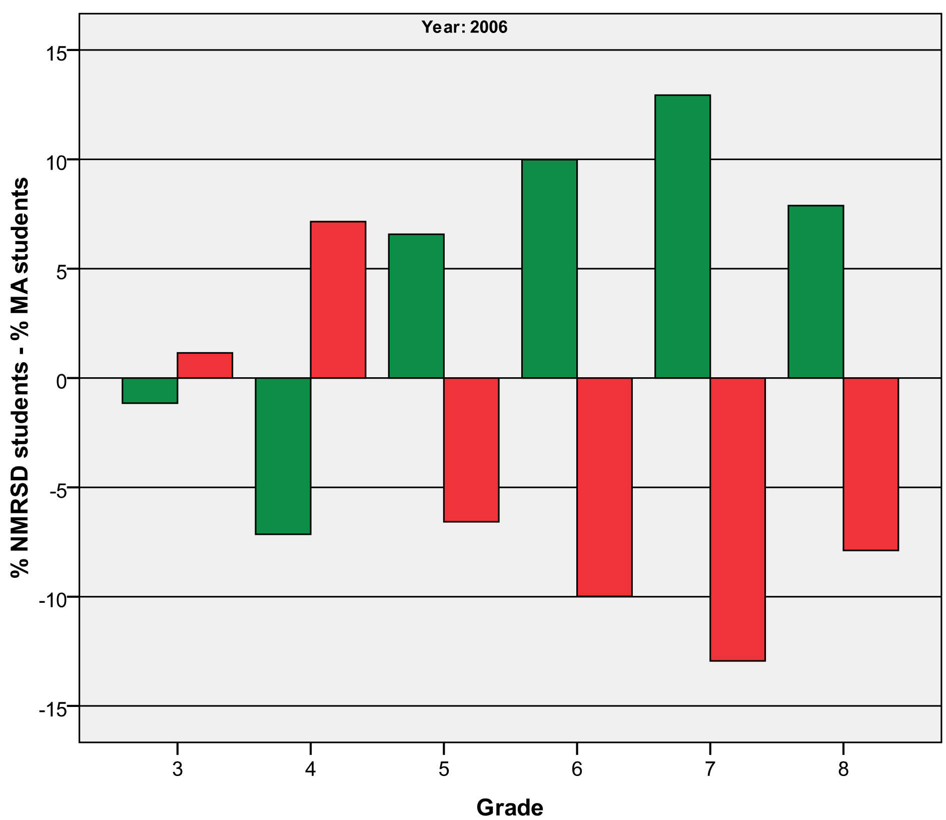

Year: 2002
20
10
0
-10
% NMRSD students - % MA students
4
6
8
10
Grade

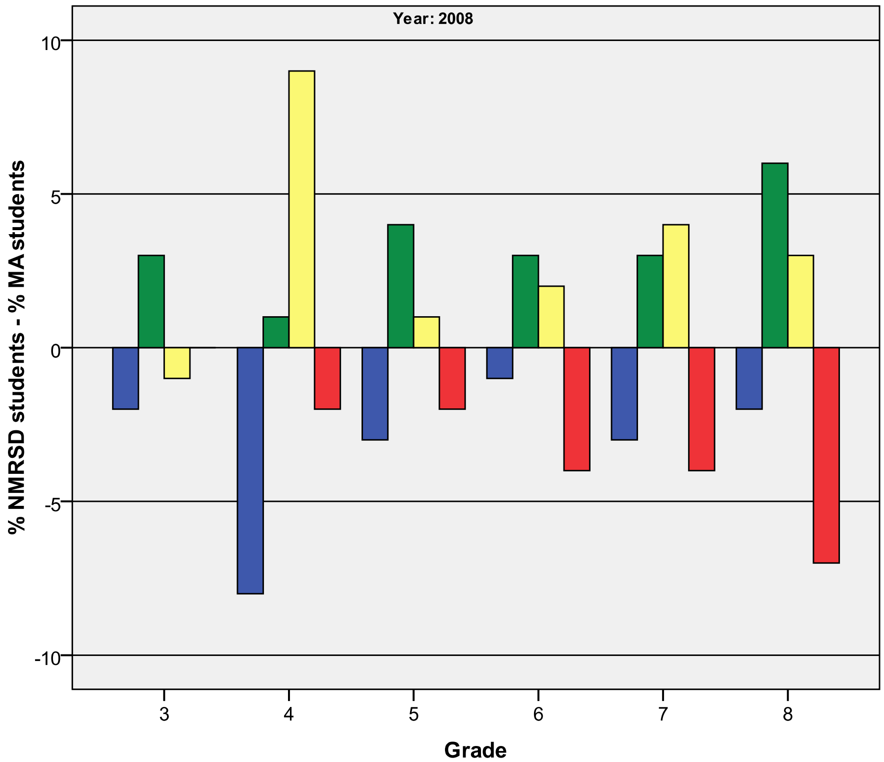

Year: 2002
20
10
0
-10
-20
% NMRSD students - % MA students
4
6
8
Grade

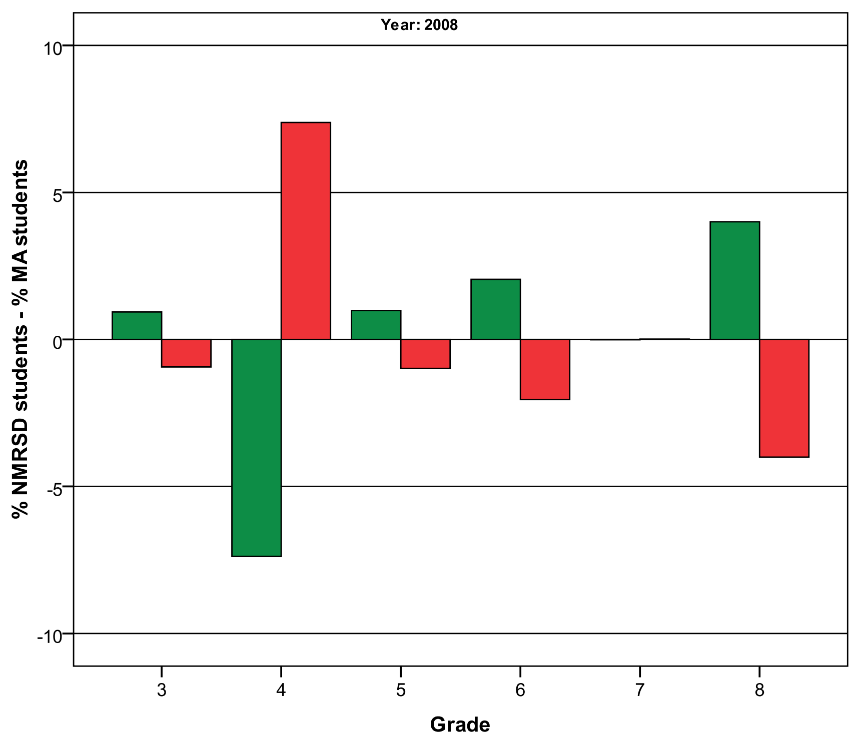

Year: 2004
% NMRSD students - % MA students
15
10
5
0
-5
-10
-15
4
6
8
10
Grade


Year: 2004
% NMRSD students - % MA students
15
10
5
0
-5
-10
-15
4
6
8
Grade


Year: 2006
% NMRSD students - % MA students
15
10
5
0
-5
-10
-15
3
4
5
6
7
8
Grade


Year: 2006
% NMRSD students - % MA students
15
10
5
0
-5
-10
-15
3
4
5
6
7
8
Grade


Year: 2008
10
5
0
-5
-10
% NMRSD students - % MA students
3
4
5
6
7
8
Grade


Year: 2008
10
5
0
-5
-10
% NMRSD students - % MA students
3
4
5
6
7
8
Grade

The 2008 graph for %NMRSD students minus %MA students, above, seems to suggest that either MA students are improving or that NMRSD students are worsening. We have seen that this is not the case when we consider mean scores, and the same is true when we look at the paired categories, Advanced plus Proficient and the Needs Improvement plus Failing.

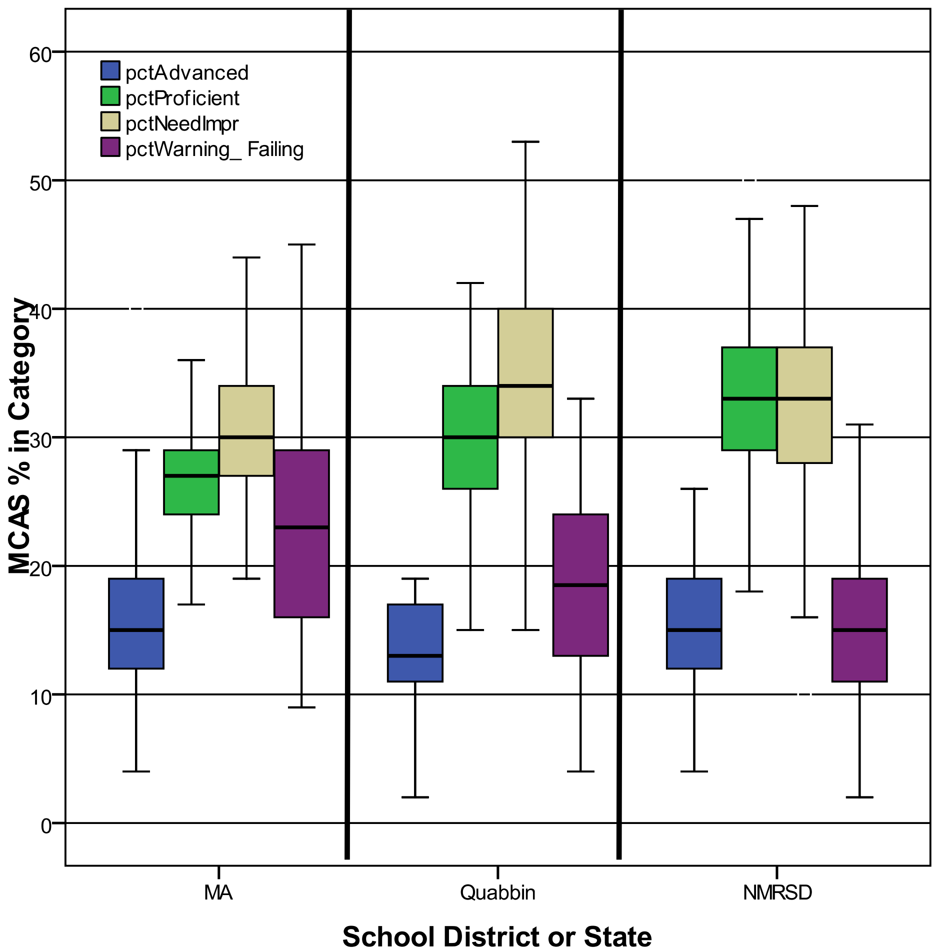


Figure 14. MA, NMRSD and Quabbin Mean MCAS Percent in Proficiency Categories, 1998-2008.

Quabbin School District is near NMRSD and similar demographically. The categorical comparison we carried out in Figure 14, above, again shows NMRSD students achieving better results. There is little to suggest that further analysis would be informative; Quabbin students continue to improve over the period, as do NMRSD and MA.

Appendix 2 – Individual Teacher Results, General Linear Model

Dependent Variable: MCAS Raw Mathematics Score

Model: (Intercept) + SM + teacher + Free Lunch + Special Ed

The coefficient indicates the direction and amplitude of the effect of the particular teacher and other model effect on MCAS scores.

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	26.321	.5646	25.215	27.428	2173.426	1	.000
[SM=0]	-1.309	.4656	-2.222	-.397	7.905	1	.005
[Tchr=10]	-1.973	1.0450	-4.021	.076	3.563	1	.059
[Tchr=16]	-3.223	1.0510	-5.283	-1.163	9.404	1	.002
[Tchr=18]	-4.607	.9441	-6.458	-2.757	23.819	1	.000
[Tchr=25]	-2.384	.6463	-3.651	-1.118	13.612	1	.000
[Tchr=32]	-3.425	1.0813	-5.545	-1.306	10.035	1	.002
[Tchr=35]	-1.321	.7586	-2.808	.166	3.034	1	.082
[Tchr=37]	-9.246	1.0875	-11.378	-7.115	72.291	1	.000
[Tchr=40]	-3.716	.7410	-5.168	-2.263	25.146	1	.000
[Tchr=42]	-3.649	.9440	-5.499	-1.798	14.938	1	.000
[Tchr=44]	-.716	.8904	-2.461	1.029	.647	1	.421
[Tchr=46]	.646	1.0191	-1.351	2.644	.402	1	.526
[Tchr=53]	-9.265	.9896	-11.205	-7.326	87.652	1	.000
[Tchr=54]	-6.413	.8757	-8.129	-4.697	53.627	1	.000
[Tchr=56]	-6.576	.9179	-8.375	-4.777	51.332	1	.000
[Tchr=58]	-4.833	.9802	-6.754	-2.912	24.308	1	.000
[Tchr=61]	-3.120	.9439	-4.970	-1.270	10.926	1	.001
[Tchr=62]	-7.531	1.0808	-9.650	-5.413	48.559	1	.000
[Tchr=65]	-1.922	1.0723	-4.024	.179	3.214	1	.073
[Tchr=67]	-6.061	.8248	-7.678	-4.444	53.994	1	.000
[Tchr=68]	-4.143	.5866	-5.293	-2.994	49.894	1	.000
[Tchr=77]	-5.740	.9261	-7.556	-3.925	38.419	1	.000
[Tchr=80]	-4.137	.9633	-6.025	-2.249	18.442	1	.000
[Tchr=84]	-2.448	.9684	-4.346	-.550	6.389	1	.011

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
[Tchr=86]	-5.259	.6394	-6.512	-4.006	67.650	1	.000
[Tchr=91]	-8.413	1.0444	-10.460	-6.366	64.887	1	.000
[Tchr=92]	-.513	1.0378	-2.548	1.521	.245	1	.621
[Tchr=93]	-2.183	1.0246	-4.192	-.175	4.541	1	.033
[Tchr=95]	-7.509	.9900	-9.450	-5.569	57.540	1	.000
[Tchr=99]	-3.278	.8829	-5.009	-1.548	13.788	1	.000
[Tchr=101]	-9.348	1.0441	-11.394	-7.301	80.150	1	.000
[Tchr=105]	-3.830	.9743	-5.740	-1.921	15.457	1	.000
[Tchr=107]	-2.909	.9631	-4.796	-1.021	9.123	1	.003
[Tchr=109]	-3.351	.5684	-4.465	-2.237	34.755	1	.000
[Tchr=113]	-9.326	1.0800	-11.443	-7.209	74.555	1	.000
[Tchr=124]	-3.550	1.0312	-5.571	-1.529	11.853	1	.001
[Tchr=126]	-2.994	.9773	-4.909	-1.078	9.384	1	.002
[Tchr=127]	-4.586	.5840	-5.731	-3.441	61.667	1	.000
[Tchr=130]	-5.062	.6732	-6.382	-3.742	56.531	1	.000
[Tchr=133]	-6.579	.6354	-7.825	-5.334	107.207	1	.000
[Tchr=135]	-4.305	.9736	-6.213	-2.397	19.554	1	.000
[Tchr=137]	-7.354	1.0375	-9.387	-5.321	50.245	1	.000
[Tchr=141]	-3.351	.5882	-4.504	-2.198	32.464	1	.000
[Tchr=148]	-5.757	.9261	-7.572	-3.942	38.646	1	.000
[Tchr=149]	-2.330	.5804	-3.467	-1.192	16.110	1	.000
Free Lunch	-3.033	.3614	2.325	3.741	70.442	1	.000
Special Education	-9.717	.3330	9.064	10.369	851.381	1	.000

Here, it may be appropriate to quote from the remarks of Dr. Mary Waight, NMRSD Associate Superintendent (retired), to the National Math Panel in 2006:

> *Improving outcomes for students in mathematics is dependent on a number of factors, chief among them a teacher with a strong math background, ongoing professional development, administrative support and involvement, and a mathematics program that encourages mathematical understanding.*[14]

Appendix 3 – χ^2 on categories on MCAS

Table 25. x² results for Categorical MCAS results, by grade-year							
Year	Grade	**N** **%**	Ad-vanced	Profi-cient	Needs Improvement	Warn-ing/ Failure	x^2
1998	4	MA	8,147	17,036	32,590	17,036	x2= 23.018 (P = <0.001)
		%	11	23	44	23	
		NMRSD	63	118	176	63	
		%	15	28	42	15	
	8	MA	5,456	15,686	17,731	28,643	x2= 23.289 (P = <0.001)
		%	8	23	26	42	
		NMRSD %	30	96	130	115	
		%	8	26	35	31	
	10	MA	4,291	10,420	14,711	31,874	x2= 21.242 (P = <0.001)
		%	7	17	24	52	
		NMRSD	24	49	92	106	
		%	9	18	34	39	
1999	4	MA	9,205	18,411	33,753	14,575	x2= 24.781 (P = <0.001)
		%	12	24	44	19	
		NMRSD	36	82	98	22	
		%	15	34	41	9	
	8	MA	4,228	15,502	21,844	28,186	x2= 22.083 (P = <0.001)
		%	6	22	31	40	
		NMRSD	19	70	84	60	
		%	8	30	36	26	
	10	MA	5,465	9,108	13,965	32,181	x2= 8.901 (P = 0.031)
		%	9	15	23	53	
		NMRSD	16	43	65	103	
		%	7	19	29	45	
2000	4	MA	9,298	21,696	32,545	13,948	x2= 32.807 (P = <0.001)
		%	12	28	42	18	
		NMRSD	48	152	160	36	
		%	12	38	40	9	
	8	MA	7,057	16,936	19,053	27,521	x2= 6.571 (P = 0.087)
		%	10	24	27	39	
		NMRSD	30	104	112	130	
		%	8	28	30	35	
	10	MA	9,099	10,919	13,345	27,298	x2= 12.324 (P = 0.006)
		%	15	18	22	45	
		NMRSD	36	58	81	103	

Table 25. x^2 results for Categorical MCAS results, by grade-year							
Year	Grade	**N** **%**	Ad-vanced	Profi-cient	Needs Improve ment	Warn-ing/ Failure	x^2
		%	13	21	29	37	
2001	4	MA	7,548	18,115	34,720	14,341	x2= 0.0032 (P = 1.000)
		%	10	24	46	19	
		NMRSD	38	92	176	73	
		%	10	24	46	19	
	6	MA	9,995	17,683	23,065	25,371	x2= 0.0056 (P = 1.000)
		%	13	23	30	33	
		NMRSD	51	91	119	130	
		%	13	23	30	33	
	8	MA	7,930	16,580	24,510	22,348	x2= 0.0042 (P = 1.000)
		%	11	23	34	31	
		NMRSD	46	97	143	131	
		%	11	23	34	31	
	10	MA	11,550	17,325	19,250	16,042	x2= 0.0000 (P = 1.000)
		%	18	27	30	25	
		NMRSD	54	81	90	75	
		%	18	27	30	25	
2002	4	MA	8,881	19,983	31,084	14,062	x2= 43.480 (P = <0.001)
		%	12	27	42	19	
		NMRSD	60	148	160	32	
		%	15	37	40	8	
	6	MA	10,396	22,391	23,191	23,991	x2= 48.032 (P = <0.001)
		%	13	28	29	30	
		NMRSD	72	152	112	64	
		%	18	38	28	16	
	8	MA	8,063	16,859	24,189	24,189	x2= 42.214 (P = <0.001)
		%	11	23	33	33	
		NMRSD	53	123	160	74	
		%	13	30	39	18	
	10	MA	12,889	15,467	19,979	16,112	x2= 25.172 (P = <0.001)
		%	20	24	31	25	
		NMRSD	72	88	74	41	
		%	26	32	27	15	
2003	4	MA	8,736	20,383	32,031	11,648	x2= 39.020 (P = <0.001)
		%	12	28	44	16	
		NMRSD	50	119	173	14	
		%	14	33	49	4	
	6	MA	12,086	19,639	24,171	19,639	
		%	16	26	32	26	
		NMRSD	68	125	140	45	
		%	18	33	37	12	
	8	MA	9,107	18,974	22,768	25,045	x2= 69.413 (P = <0.001)

Table 25. x^2 results for Categorical MCAS results, by grade-year							
Year	Grade	**N** **%**	Ad-vanced	Profi-cient	Needs Improve ment	Warn-ing/ Failure	x^2
		%	12	25	30	33	
		NMRSD	52	153	137	60	
		%	13	38	34	15	
	10	MA	15,746	17,714	19,026	13,122	x2= 12.460 (P = 0.006)
		%	24	27	29	20	
		NMRSD	78	107	98	42	
		%	24	33	30	13	
2004	4	MA	9,875	19,751	31,037	9,875	x2= 25.913 (P = <0.001)
		%	14	28	44	14	
		NMRSD	61	140	152	27	
		%	16	37	40	7	
	6	MA	12,597	19,266	23,712	18,525	x2= 45.032 (P = <0.001)
		%	17	26	32	25	
		NMRSD	80	155	130	54	
		%	19	37	31	13	
	8	MA	9,864	19,728	24,281	22,004	x2= 40.946 (P = <0.001)
		%	13	26	32	29	
		NMRSD	59	137	141	59	
		%	15	35	36	15	
	10	MA	19,335	18,668	18,668	10,001	x2= 38.403 (P = <0.001)
		%	29	28	28	15	
		NMRSD	126	109	68	19	
		%	39	34	21	6	
2005	4	MA	10,065	18,693	31,634	10,784	x2= 4.156 (P = 0.245)
		%	14	26	44	15	
		NMRSD	49	105	154	42	
		%	14	30	44	12	
	6	MA	12,545	21,400	22,138	16,972	x2= 45.642 (P = <0.001)
		%	17	29	30	23	
		NMRSD	92	135	99	40	
		%	25	37	27	11	
	8	MA	9,851	19,703	22,734	23,492	x2= 32.115 (P = <0.001)
		%	13	26	30	31	
		NMRSD	42	125	141	72	
		%	11	33	37	19	
	10	MA	22,860	18,153	16,136	10,085	x2= 87.862 (P = <0.001)
		%	34	27	24	15	
		NMRSD	162	87	29	12	
		%	56	30	10	4	
2006	3	MA	2,792	33,500	22,333	11,167	x2= 0.766 (P = 0.857)
		%	4	48	32	16	
		NMRSD	14	165	120	53	
		%	0	47	34	15	

Table 25. x^2 results for Categorical MCAS results, by grade-year

Year	Grade	**N** **%**	Ad-vanced	Profi-cient	Needs Improve ment	Warn-ing/ Failure	x^2
	4	MA	10,545	17,575	31,635	10,545	x2= 16.158 (P = 0.001)
		%	15	25	45	15	
		NMRSD	31	82	187	44	
		%	9	24	54	13	
	5	MA	12,196	18,653	24,393	16,501	x2= 10.759 (P = 0.013)
		%	17	26	34	23	
		NMRSD	68	107	121	57	
		%	19	30	34	16	
	6	MA	12,322	21,020	21,020	18,121	x2= 34.370 (P = <0.001)
		%	17	29	29	25	
		NMRSD	58	157	111	58	
		%	15	41	29	15	
	7	MA	8,839	20,625	24,307	20,625	x2= 35.045 (P = <0.001)
		%	12	28	33	28	
		NMRSD	43	143	107	61	
		%	12	40	30	17	
	8	MA	9,036	21,084	23,343	21,837	x2= 22.842 (P = <0.001)
		%	12	28	31	29	
		NMRSD	55	149	141	81	
		%	13	35	33	19	
	10	MA	28,926	19,525	15,186	8,678	x2= 38.128 (P = <0.001)
		%	40	27	21	12	
		NMRSD	165	81	39	15	
		%	55	27	13	5	
	3	MA	13,378	28,869	16,899	11,266	x2= 15.811 (P = 0.001)
		%	19	41	24	16	
		NMRSD	66	174	73	35	
		%	19	50	21	10	
	4	MA	13,219	20,177	27,134	9,045	x2= 16.787 (P = <0.001)
		%	19	29	39	13	
		NMRSD	39	104	161	54	
		%	11	29	45	15	
2007	5	MA	13,372	22,521	21,817	12,668	x2= 16.322 (P = <0.001)
		%	19	32	31	18	
		NMRSD	48	136	113	44	
		%	14	40	33	13	
	6	MA	14,380	23,008	20,132	14,380	x2= 12.987 (P = 0.005)
		%	20	32	28	20	
		NMRSD	64	146	86	61	
		%	18	41	24	17	
	7	MA	10,904	22,535	21,808	17,447	x2= 13.496 (P = 0.004)
		%	15	31	30	24	
		NMRSD	53	140	125	64	

Table 25. x^2 results for Categorical MCAS results, by grade-year							
Year	Grade	**N** **%**	Ad-vanced	Profi-cient	Needs Improve ment	Warn-ing/ Failure	x^2
		%	14	37	33	17	
	8	MA	12,489	20,570	22,040	18,367	x2= 21.563 (P = <0.001)
		%	17	28	30	25	
		NMRSD	80	127	98	62	
		%	22	35	27	17	
	10	MA	29,255	19,265	15,698	6,422	x2= 19.031 (P = <0.001)
		%	41	27	22	9	
		NMRSD	145	91	47	12	
		%	49	31	16	4	
2008	3	MA	17,599	25,343	17,599	9,856	x2= 1.450 (P = 0.694)
		%	25	36	25	14	
		NMRSD	76	129	80	46	
		%	23	39	24	14	
	4	MA	14,291	20,722	27,153	9,289	x2= 20.603 (P = <0.001)
		%	20	29	38	13	
		NMRSD	41	103	164	38	
		%	12	30	47	11	
	5	MA	15,565	21,224	21,224	12,027	x2= 3.901 (P = 0.272)
		%	22	30	30	17	
		NMRSD	66	117	107	52	
		%	19	34	31	15	
	6	MA	16,486	23,654	18,636	12,902	x2= 4.435 (P = 0.218)
		%	23	33	26	18	
		NMRSD	77	125	97	49	
		%	22	36	28	14	
	7	MA	10,975	23,413	21,218	17,560	x2= 7.108 (P = 0.069)
		%	15	32	29	24	
		NMRSD	42	122	115	70	
		%	12	35	33	20	
	8	MA	13,940	22,010	19,809	17,608	x2= 14.588 (P = 0.002)
		%	19	30	27	24	
		NMRSD	65	138	115	65	
		%	17	36	30	17	
	10	MA	30,606	20,641	13,523	6,406	x2= 32.618 (P = <0.001)
		%	43	29	19	9	
		NMRSD	168	92	40	6	
		%	55	30	13	2	