



UpFront
Consulting

Math Corps Evaluation

*2009-10 Findings Report
in St. Cloud Area School District 742*

Prepared by:
UpFront Consulting
September 9, 2010



CONTENTS

Key points summary	3
Program overview	4
The tutoring process	4
Student profile	5
Overall outcomes	7
MCA II findings	7
MAP scores	8
Progression level	9
Changing the level needed to progress	9
Data on progression	9
Tutoring time	11
Minutes tutored, weeks participating	11
Success of students based on average minutes tutored	13
Comparison to 2008-09	15
Fidelity	17
Impact of fidelity checks in 2009-10	17
Comparison between 2008-09 and 2009-10	18
Factors that influence achievement	20
Ethnicity	20
Socio-economic background	21
English language learners	22
Impact of second-year tutors	23
Second-year tutors	23
Students of second-year tutors	24

KEY POINTS SUMMARY

UpFront Consulting conducted an evaluation of the Math Corps program in St. Cloud Area schools. The evaluation analyzed data collected during the 2009-10 school year, and compared it to data collected during a previous evaluation in 2008-09. Key findings are as follows:

- Overall, participating students in 2009-10 outperformed students in 2008-09. This was true for two measures—scores on the Minnesota Comprehensive Assessment (MCA II/MTELL) and the Measures of Academic Progress (MAP). Students in grades 6 through 8 had a higher percentage of students who made the MCA proficiency rating than grades 4 and 5. However, in MAP scores (which is probably a more robust measure of progress because it takes into account the different starting points of each student) grades four and six did best.
- This year the program raised the level students must achieve on the unit assessments to progress to the next level. Overall, students did better this year. Any slowing of progress through the units caused by the raised threshold seems to be offset by the increased content mastery.
- This year the program also increased the minutes per week students are tutored. The data suggests that there is a small improvement when grades 4 to 6 students get 70 minutes per week instead of 60. For grades 7 and 8 the improvement is more dramatic, both at the under/over 70 minutes and under/over 80 minutes levels. The data suggests that as the program gets closer to meeting the tutoring time target (90 minutes per week), students do better.
- The data does not predict which factor—placing the advancement threshold at 85% or increasing tutoring time to 90 minutes per week—is more important in boosting achievement.
- The increased rigor in the fidelity process introduced in 2009-10 may also have contributed to the program's greater success. However, there were so few instances of tutors not administering the program with fidelity in 2009-10 that the data is inconclusive.
- The program does equally well with students of color as with Caucasian students. The program also does well with low socio-economic status students, who scored within 2 percentage points of their better-off peers on both MCA-II and MAP. Limited English Proficiency students showed slightly more progress than students who have a better command of English.
- Students in schools rated as lower in supporting the program (based on the survey all 22 tutors) did not perform any differently than students in high-support schools.
- The data on students of second-year tutors was inconclusive.

More information about each of these points is in the body of the report.

PROGRAM OVERVIEW

Minnesota Math Corps began working in four elementary and one junior high in St. Cloud Area School District 742 in 2008-09. The program was expanded in 2009-10 to include all eight elementary schools, both junior highs, and the junior high students at the Area Learning Center and at Kennedy (Kennedy Community School is a K-8 building).

Math Corps works with students in grades four through eight who scored just below proficiency on the state's No Child Left Behind (NCLB) exam, the Minnesota Comprehensive Assessment II (MCA-II) or the Minnesota Test of English Language Learners (MTELL). Two other criteria for student selection are 1) the student is not receiving special education services, and 2) the student's teacher has recommended the student for participation.

In 2008-09, the five District 742 buildings were served by twelve tutors, supported by one Master Coach. In 2009-10, the 11 buildings were served by 22 tutors and the Master Coach.

In 2008-09, 160 students participated in the program. In 2009-10, there were 256 participants.

The tutoring process

In 2008-09, the tutors provided between 60 and 90 minutes of one-on-one tutoring to each participating student each week. In 2009-10, the program set a minimum of 90 minutes per week per elementary student; junior high students were generally scheduled for more minutes per week.

Tutoring sessions took place during the school day, outside of the classroom, and at times that did not conflict with regular Language Arts or Mathematics instruction.

Tutoring was individualized to meet student needs identified by the previous MCA-II (MTELL) test substrands and by the strands of the Measures of Academic Progress (MAP) Mathematics test, produced by Northwest Evaluation Association (NWEA).

Members began the tutoring process with each student by administering a diagnostic test. Results of the test were used to determine which lessons would be taught.

After each lesson (generally from one to three tutoring sessions) students would complete a related technology activity. If the student scored poorly on the activity, the Member would complete support activities and redo the original activity. If the student scored poorly the second time through the original technology activity, the Member would repeat the lesson.

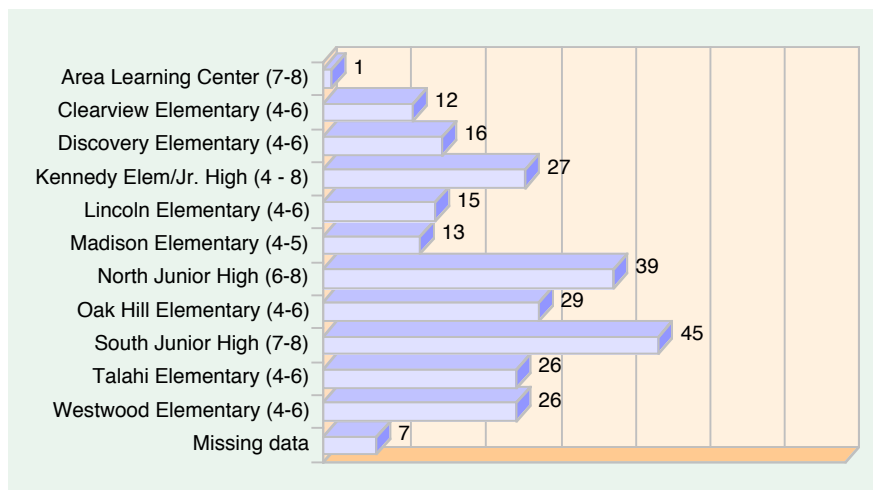
After completing all the selected lessons of the unit and the corresponding technology activities, the student would complete a unit test. In 2008-09, if the student scored less than 75% on the unit test, the Member would repeat the lesson. In 2009-10, this was raised to 85%; if the student

scored lower than the lesson was repeated.

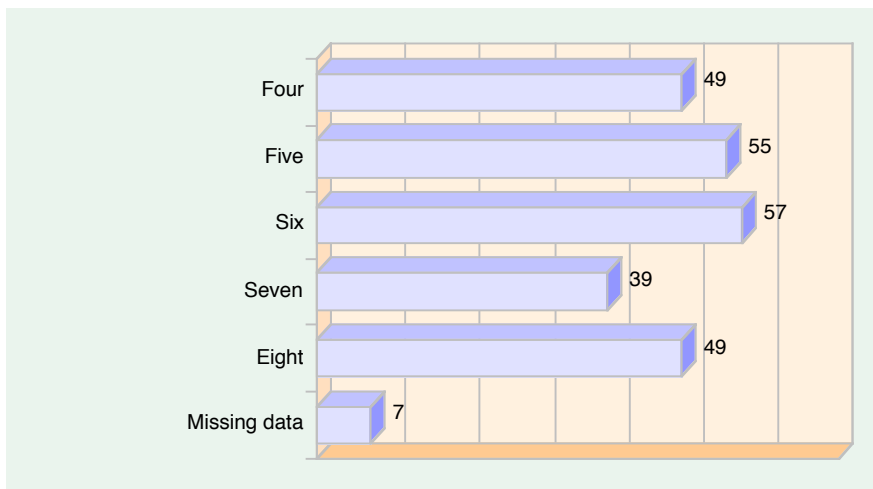
If the student scored above the threshold, the member would administer the diagnostic test for the next unit in the area of identified need. The member would repeat this process of diagnostic test, unit lessons and technology activities, then unit test, until all units were completed.

Student profile

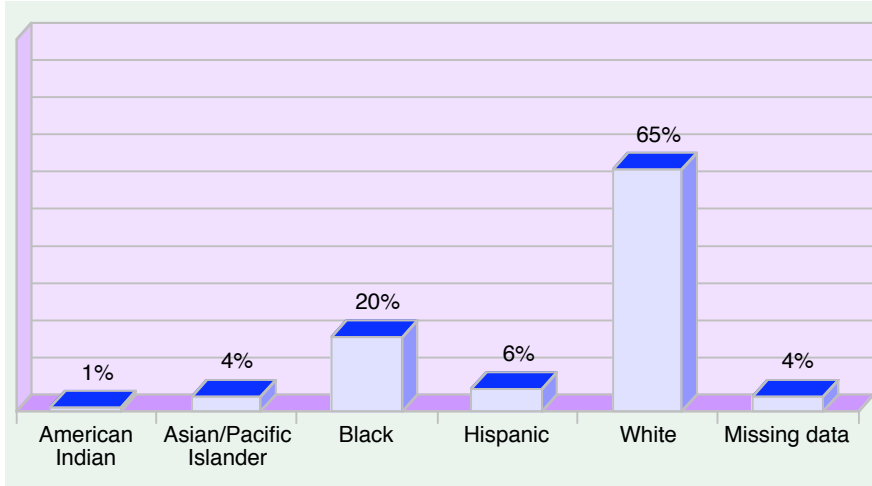
Demographic characteristics of the 256 students in the program in 2009-10 are shown in the charts below and on the subsequent pages.



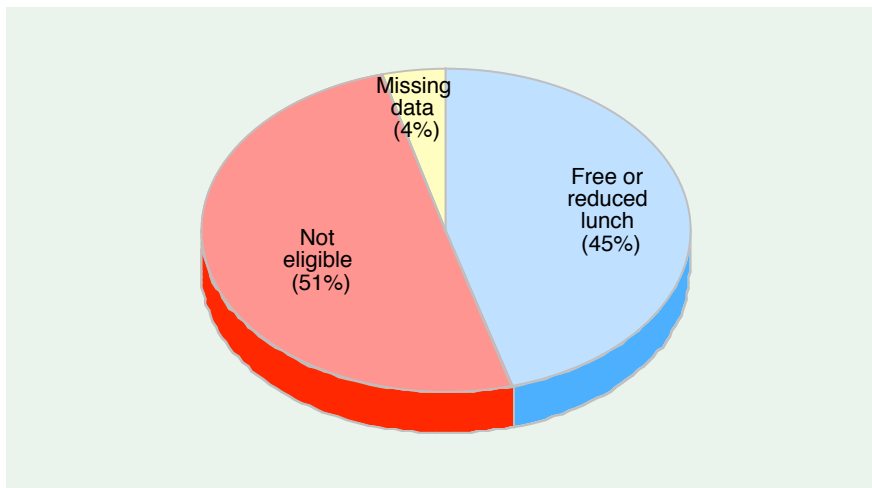
Number of students from each school served by the program. Note that this reflects the 2009-10 configuration with Kennedy serving grades K - 8, Madison K - 5, and North Junior High grades 6 - 8. ALC has students from grades 7 - 12. South Junior High serves grades 7 and 8; the remaining elementary schools were all grades K - 6.



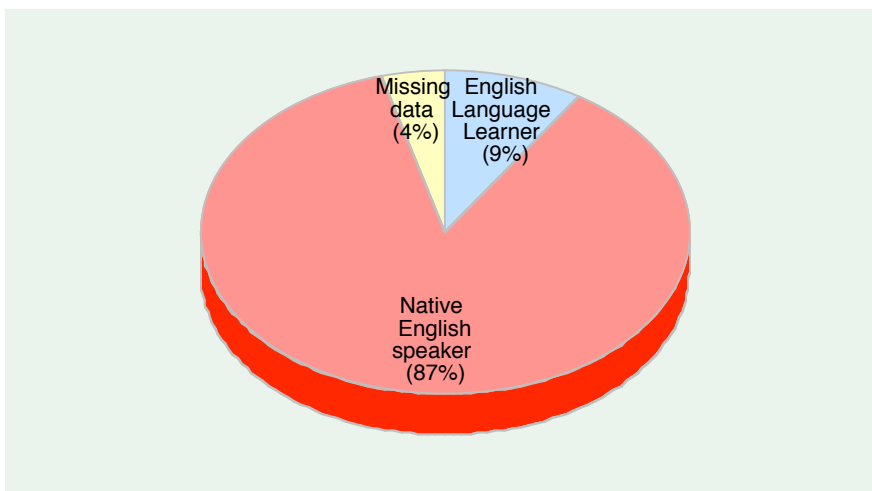
Number of students at each grade level served by the program in 2009-10.



More than six in ten students tutored by Math Corps members are White; about two in ten are Black. Overall, District 742 is 76% White and 15% Black.



Shown is the percent of participating students eligible for free or reduced lunch (FRL). FRL is an indication of family poverty. Overall, 49% of students in District 742 are eligible for free or reduced lunch.



Shown is the percent of students served by the program who have Limited English Proficiency (LEP). Overall, 10% of students in the District are classified as LEP.

OVERALL OUTCOMES IN 2009-10

The charts and tables in this section provide some insight into overall program outcomes in 2009-10. This data is presented here as background to the research questions and findings presented later in this report.

MCA-II findings

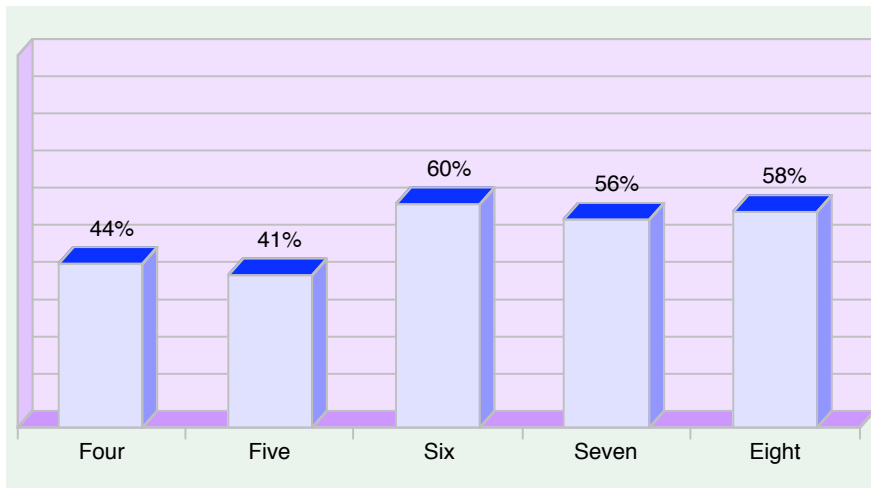
As noted earlier, students take the MCA-II exam late in each school year. The student's proficiency level is then used to determine eligibility for the Math Corps program. Students who score between 45 and 49 are eligible, (scores of 40 to 49 are considered "Partially Proficient" in MCA-II).

One measure of the program's overall success is the percent of students who achieve proficiency (a score of 50 or higher) at the end of the academic year.

The chart and table below show the results for all Math Corps students who took the MCA-II exams in the spring of 2010. Since all students would have scored from 45 to 49 the previous year, the percent of those who score 50 or higher (proficient) is one measure of program success.

All grades showed progress; the three higher grades had the highest percent of students who moved into proficiency.

Note that the MCA-II tests are progressively more difficult each year.



Percent of Math Corps students scoring 50 or above (proficient) on the MCA-II assessments in the spring of 2010, by grade level.

MCA-II, proficiency of Math Corps students, 2009-10

<u>Grade</u>	<u>n</u>	<u>Percent proficient (50 or higher)</u>	<u>Partially proficient (40 to 49)</u>
Four43	44%	42%
Five51	41%	49%
Six52	60%	37%
Seven36	56%	39%
Eight48	58%	42%

Percent of Math Corps students scoring 50 or above (proficient) on the MCA-II assessments in the spring of 2010, by grade level. Also shown is the percent who scored between 40 and 49 (partially proficient).

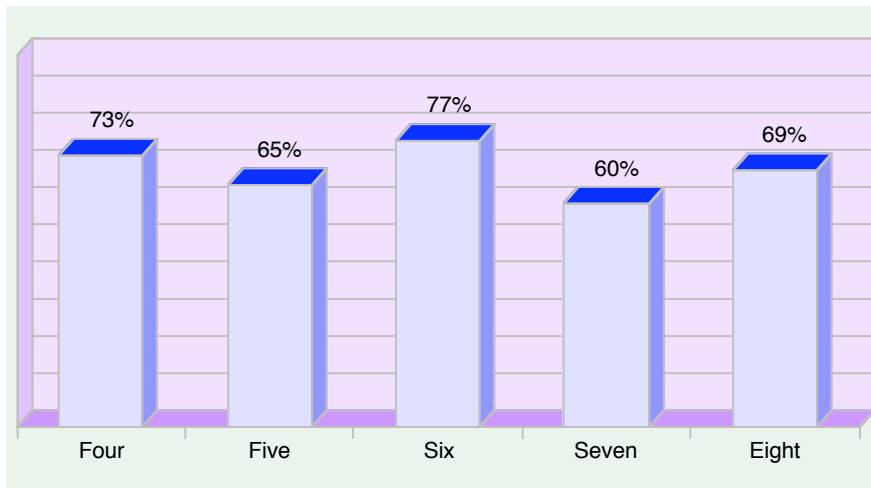
MAP scores

The chart and table show Math Corps student performance on Measures of Academic Progress (MAP) mathematics test, produced by Northwest Evaluation Association (NWEA).

Students take the MAP test three times per year, in fall, winter and spring. Based on nationwide data, NWEA provides growth targets or norms for each student. The data described in this report is the student's actual progress between fall and spring assessments compared to the progress he or she is expected to achieve, based on the nationwide norms. Note that the norms vary by grade level and by starting (fall) score. This analysis generally reports the percent of students who equaled or bested their growth target. It also examines the mean gain between fall and spring tests.

A high percentage of Math Corps students met or exceeded the growth target for their beginning score and grade level, as shown below.

Because the MAP scores are normed based on the student's starting point (the fall test), they are perhaps a slightly more robust measure of success than the MCA data shown on the previous page.



Percent of Math Corps students who met or exceeded the growth target for their grade level and starting score. Data includes only those students who took both the fall and spring MAP tests.

Percent who met or exceeded their MAP growth target

<u>Grade</u>	<u>n</u>	<u>Percent at or above norm</u>
Four44	73%
Five52	65%
Six53	77%
Seven37	60%
Eight48	69%

Number of Math Corps students by grade and percent who met or exceeded the growth target for their grade level and starting score. Data includes only those students who took both the fall and spring MAP tests

PROGRESSION LEVEL

Research question— In 2008-09, students needed to score 75% on the assessment tool to progress to the next level. In 2009-10 the bar was set at 85%. Is there a significant difference in student success between students who were progressed at 75% and students who were progressed at 85%?

Changing the level needed to progress

The change from 75% to 85% as the threshold needed for students to progress to the next level was an important one. In a conversation with the Master Coach in July, 2010, she reported verbal feedback from the tutors that it was a positive change, especially working in conjunction with the change to a target of 90 minutes of tutoring per week.

The coach believes that improvements from raising the bar to 85% will come primarily because of time. She noted that in her own observations of sessions, the overall tutoring process was similar between the two years, but that more time was spent on each unit.

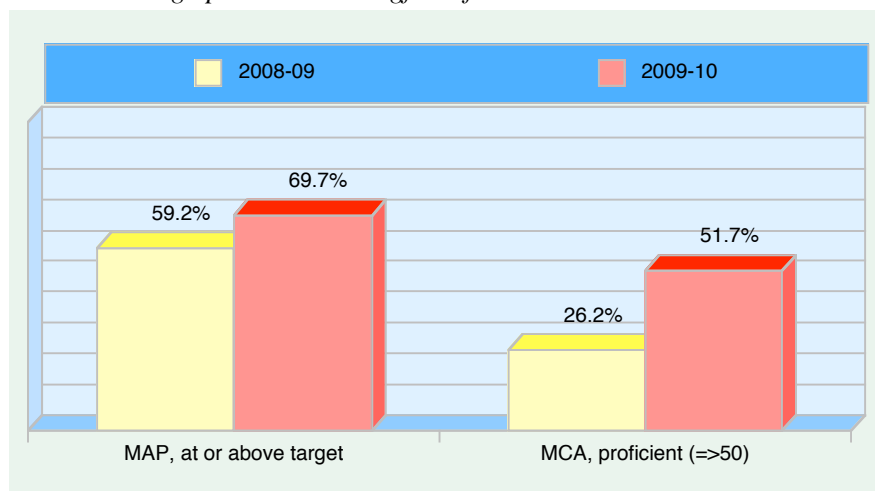
One unintended consequence was noted—the Master Coach worked over the summer to redo the unit assessments to fit the new process better.

Data on progression

The chart below shows the overall changes in achievement for the two years. By both measures, students in 2009-10 did better than students in 2008-09.

Other information about the impact of the unit assessments is unclear. For example, in the 2008-09 evaluation the number of unit assessments completed or the average score on the unit assessments did not correlate with higher achievement. Quoting from the 2008-09 evaluation report:

It does not appear that the number of unit tests completed or the unit test average provide meaningful information related to student



Performance on MAP and MCA tests, comparing students from 2008-09 and 2009-10. Students in 2009-10 had a higher threshold for progression to the next unit (85% compared to 75%) on the unit assessment. The Pearson Chi-Square Test suggests that the differences shown in both MAP and MCA are statistically significant (MAP $p = .038$; MCA-II $p = .000$).

progress toward their math goal. Neither performance indicator is related to growth on the MAP test.

In this evaluation, whether or not a student scored 85 or higher on the first assessment (and so would have gone on to the next unit rather than having additional tutoring on the first unit), also did not impact either MAP or MCA scores. For example, the percent who met or exceeded their MAP growth target was 68% for those who had to retake their first unit, and 72% for those who passed the unit on their first try. This small difference is not significant ($p = .648$). This suggests that slowing student progress through units because of the 85% threshold does not have much impact on overall achievement during the year, at least on the two large measures.

The data does not suggest whether placing the advancement threshold at 85% or increasing tutoring time to 90 minutes per week is more important in boosting achievement. The next section discusses the change in tutoring time.

Note that students are placed in their initial unit based on a preliminary assessment, so not all start with the most basic level unit.

TUTORING TIME RECEIVED BY STUDENTS

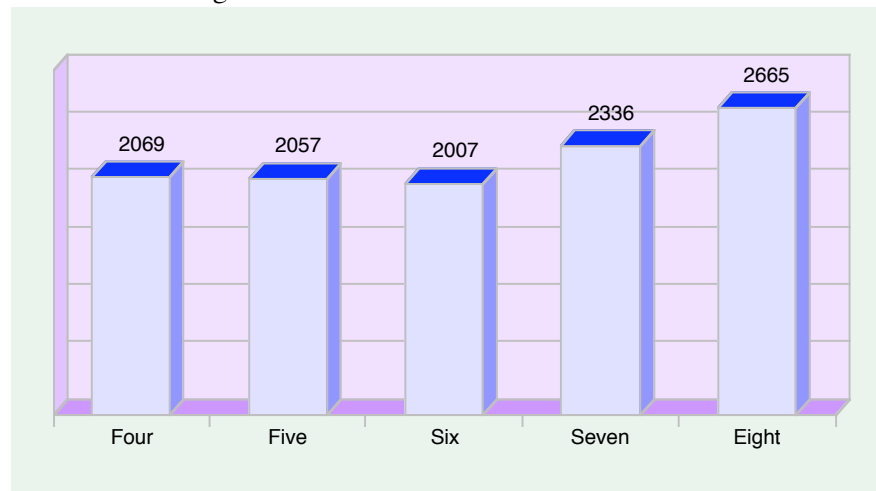
Guiding research questions—What is the proper dosage for mentoring to achieve best results? What percent of students receive 90 minutes or more of tutoring per week? Do these students do significantly better than students who receive less than 90 minutes?

Minutes tutored, weeks participating

The first step in this analysis was to examine the range of tutoring minutes experienced by students. The total number of minutes tutored during the year varied widely, from 54 minutes to 4,418. The number of weeks students participated in the program also varied widely, from one week to 36 weeks.

Overall, the following analysis uses all students who participated in the program for 12 weeks or more. This eliminated 24 students from the overall pool, leaving 226 students in the analysis.

The chart below shows the average number of minutes tutored (over the entire year) for students in each grade level who participated for 12 weeks or more. The table shows the average number of minutes students were tutored, the time they were scheduled (total time scheduled per week multiplied by the number of weeks participating in the program), and the percent of scheduled time students were actually tutored. The mean for percent of scheduled time actually tutored ranges from 70% to 73% for the five grades.



	<u>Minutes tutored</u>	<u>Minutes scheduled</u>	<u>Percent of scheduled actually completed</u>
Four	2069	2787	73%
Five	2057	2950	70%
Six	2007	2849	70%
Seven	2336	3257	71%
Eight	2665	3687	72%

“Total weeks” was defined as the period from the first week the tutor saw the student through the last week. In some cases there was a week or two in the middle when a student was not tutored. due to illness, class conflicts, etc., but for most students the weeks tutored period was continuous.

Average number of minutes students were tutored in 2009-10, by grade level. Only those who participated in the program for 12 weeks or more are included (226 total students).

The table shows the mean or average for each grade level. Each student was scheduled for a specific time each week; the total number of minutes scheduled was generally 90 minutes for grades 4 through 6, 120 minutes for grade 7 and 114 minutes for grade 8.

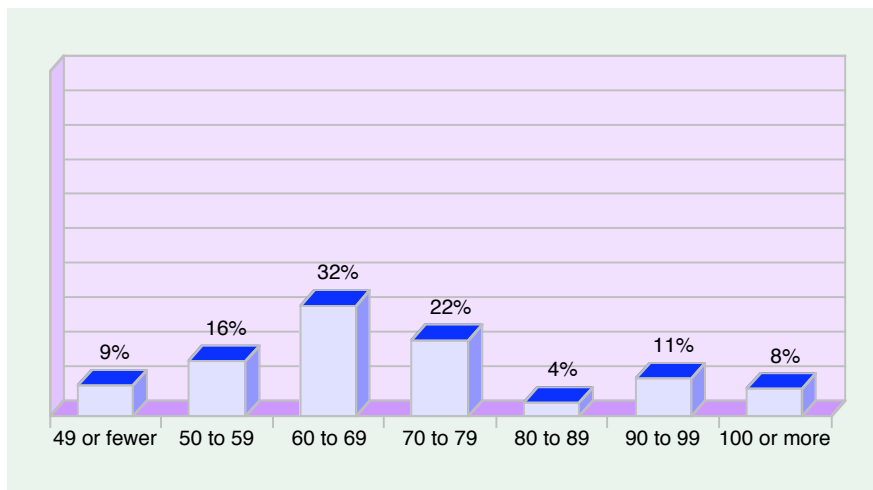
The chart below shows the percent of students who received different levels per week of tutoring. The table then breaks that down by grade level.

Note that very few of the 4th, 5th and 6th graders averaged more than 79 minutes of tutoring per week. The median (middle) is in the 60 to 69 category for all three grades. In contrast, 52% of 7th graders and 49% of 8th graders received 90 minutes or more of tutoring per week on average. The median for 7th grade is in the 90 to 99 hours per week category, for 8th graders it is in the 80 to 89 hours per week category.

This suggests that further analysis must look at the two groups of grade levels (4 through 6, 7 and 8) separately.

Based on the median responses, we examined grades 4 through 6 with 60 minutes as a threshold and 70 minutes as a threshold—students who averaged less than 60 minutes during their weeks of tutoring were compared to students who had 60 minutes or more; those with less than 70 were compared to those with 70 or more.

For grades 7 and 8, based on the different median, we used 70 minutes and 80 minutes as thresholds. That analysis begins on the next page for grades 4 to 6 and continues on the following page for grades 7 and 8.



Minutes per week students were tutored in 2009-10, in categories. Only those who participated in the program for 12 weeks or more are included (226 total students).

	<u>Four</u>	<u>Five</u>	<u>Six</u>	<u>Seven</u>	<u>Eight</u>
49 or fewer	10%	4%	11%	6%	12%
50 to 59	13%	27%	21%	3%	10%
60 to 69	35%	44%	45%	17%	10%
70 to 79	43%	25%	19%	8%	12%
80 to 89	0%	0%	2%	14%	7%
90 to 99	0%	0%	0%	33%	27%
100 or more	0%	0%	2%	19%	22%

The tables shows the minutes of tutoring per week broken down by grade level. The median for each grade is shown in bold. Students in grades 4, 5, and 6 were generally scheduled to received 90 minutes per week. Students in grade 7 were generally scheduled to receive 120 minutes per week; in grade 8 students were generally scheduled for 114 minutes per week.

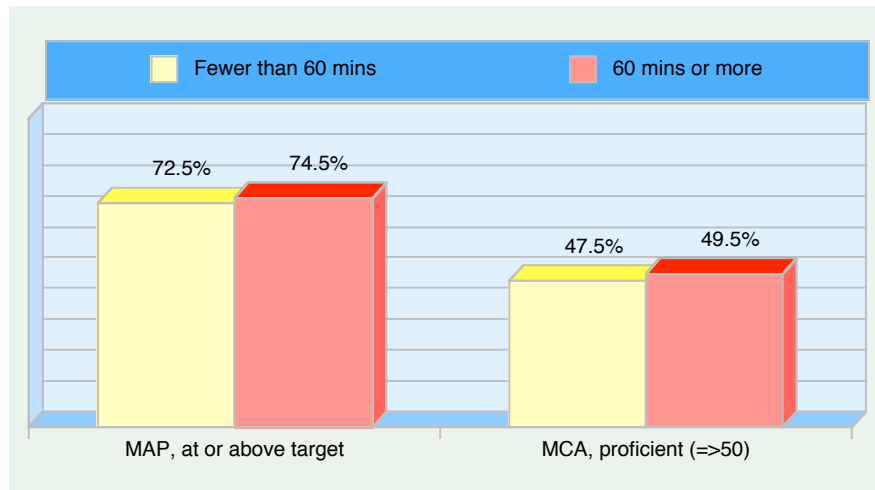
Success of students based on average minutes tutored

The two charts below show grades 4 through 6. There were too few students who receive 80 or more minutes of tutoring at these three grade levels to determine if increasing tutoring to that level would increase scores significantly.

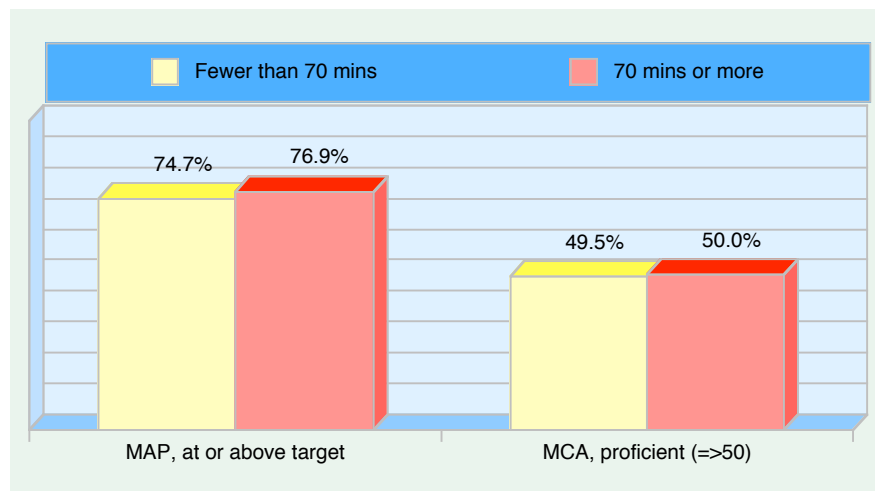
On the MAP tests, an analysis of the mean gain scores (Fall to Spring; above or or below expected growth) shows little difference at 60 minutes. However, students who received 70 minutes or more of tutoring per week averaged 5.5 points above their expected growth, compared to only 3.5 for those receiving less than 70 minutes. The difference does not meet the test of statistical significance but points in the right direction.

Overall, in grades 4 to 6, the number of minutes of tutoring may have a small impact on outcomes, but other factors likely have an influence as well.

Analysis of the mean gain above or below expected growth used a t-test to look at the means of groups with different amounts of tutoring. The difference between those with 70 or more minutes and those with less than 70 has a probability of .214.



Grades four through six, students who averaged fewer than 60 minutes of tutoring per week compared to those who averaged 60 or more. The small differences are not statistically significant. The n for these two groups is 40 for the fewer than 60 group and 98 for the 60 or greater group.



Grades four through six, students who averaged fewer than 70 minutes of tutoring per week compared to those who averaged 70 or more. The small differences are not statistically significant. The n for these two groups is 93 for the fewer than 70 group and 38 for the 70 or greater group.

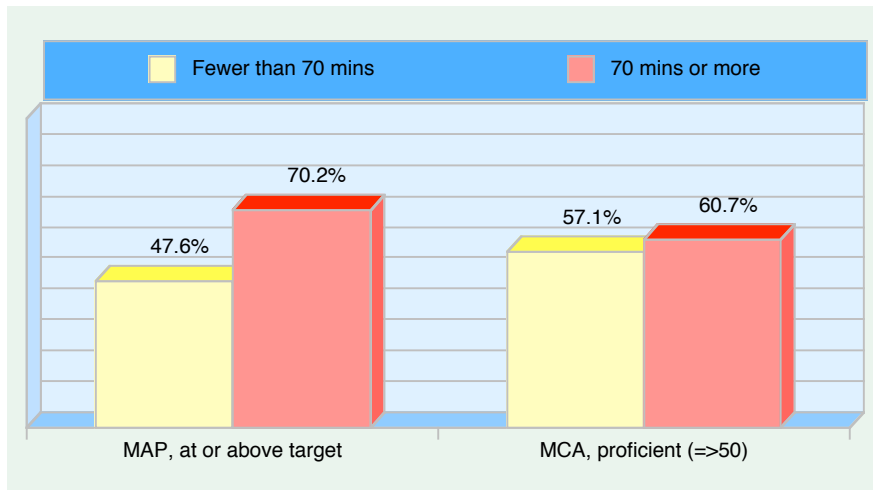


The data is much clearer in grades 7 and 8, as shown in the charts below. A much higher percent of 7th and 8th graders who were tutored 70 minutes or more per week were at or above their MAP growth target compared to those receiving less tutoring. The difference is very close to statistical significance at the 95% confidence interval.

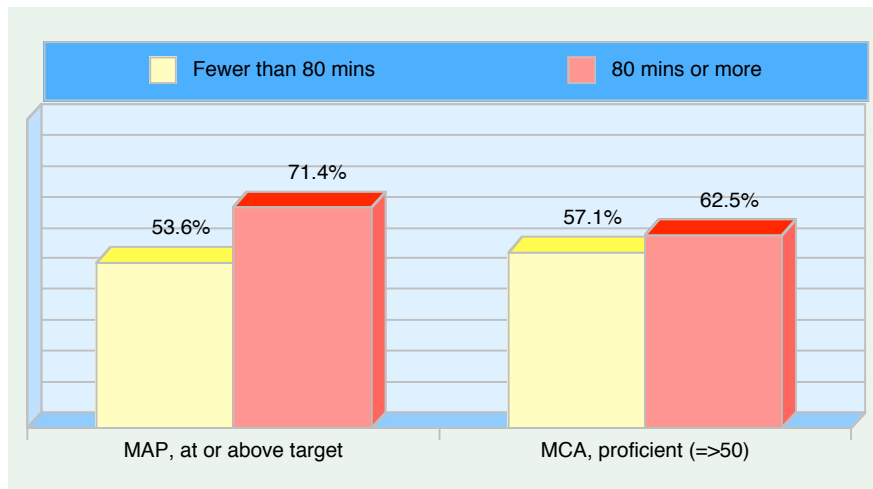
The same relationship was found using 80 minutes as the threshold, although the difference is slightly less pronounced.

Further, an analysis of the mean amount above or below the expected growth in MAP fall-to-spring scores shows that 7th and 8th grades students tutored 70 minutes or more did better (had a higher average above expected growth) than their peers with less tutoring. The differences just missed being statistically significant. The effect continues with the threshold at 80; those who averaged 80 minutes or more did better than those who received less than 80, again just missing significance.

Analysis of the mean gain above or below expected growth used a t-test to look at the averages of groups with different amounts of tutoring. The difference between those with 70 or more minutes and those with less than 70 has a probability of .095. With the threshold at 80 minutes, the probability is even smaller at .070. Both are just above the level of significance ($p < .05$).



Grades seven through nine, students who averaged fewer than 70 minutes of tutoring per week compared to those who averaged 70 or more. The small differences in the percent who were at or above their targeted MAP gain approaches statistical significance ($p = .065$). The n for these two groups is 21 for the fewer than 70 group and 56 for the 70 or greater group.



Grades seven through nine, students who averaged fewer than 80 minutes of tutoring per week compared to those who averaged 80 or more. The small differences are not statistically significant ($p = .114$). The n for these two groups is 28 for the fewer than 80 group and 49 for the 80 or greater group.

Comparison to 2008-09

The evaluation conducted in 2008-09 examined the relationship between hours tutored and achievement, by separating participating students into two groups (full treatment and partial treatment; with the threshold between the groups being 50% of intended hours of tutoring). Quoting from that report:

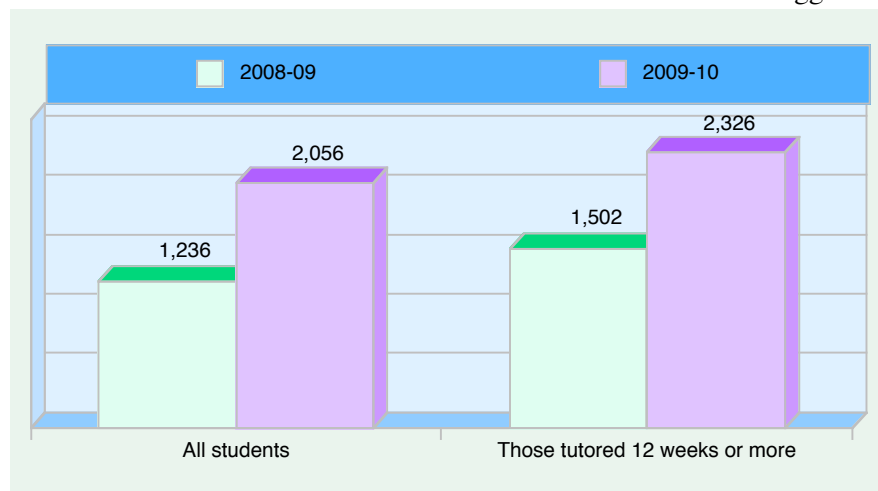
Several analyses were run to determine differences between groups, including t-tests and chi-square. None of the differences between groups were statistically significant. The difference in Fall/Spring PEG (percent of expected growth) between the Full Treatment group (155.87%) and the Comparison group (118.55%) approached significance ($t = -1.866, p = .064$), with students in the Full Treatment group demonstrating a slightly higher percent of expected growth on the MAP test from Fall-to-Spring than the students in the comparison group. However, the percent of students who were proficient on the MCA-II (or MTELL) was virtually the same across groups.

The comparison group described in the 2008-09 report was a group of students who met the criteria for the program, but did not participate.

Note that this analysis mirrors earlier findings for 2009-10, suggesting that the MAP test is a more robust measure of student achievement than the MCA scores.

A comparison between the two years of all students who received 12 weeks or more of tutoring yields some interesting findings. First, as shown in the chart below, students on average did receive more hours of tutoring in 2009-10 compared to 2008-09. Students who were in the program for 12 weeks or more in 2009-10 on averaged received 13% more tutoring minutes than similar students in 2008-09.

The analysis then compared the achievement of these students by year. The chart on the next page shows the results. This suggests that students who received 12 or more weeks of tutoring benefitted from the increase in the number of minutes tutored in 2008-09. The data does not suggest



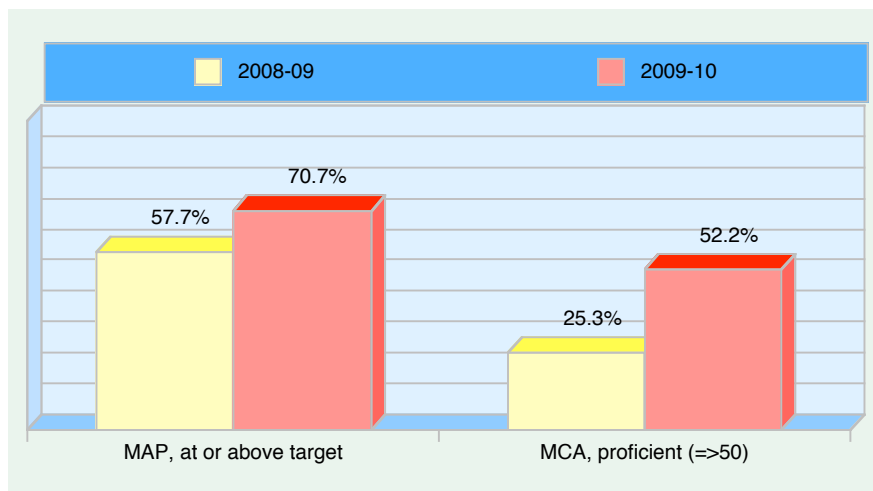
Average minutes of tutoring received by all students, and by students with approximately 12 weeks or more of tutoring, comparing 2008-09 with 2009-10. Note that the 2008-09 data is approximate while the 2009-10 data is more accurate, since the data included the number of weeks in the program.

whether the increase is due more to the increase in the number of minutes tutored, or the changed threshold for moving on to the next unit (from 75% to 85% on the unit assessment). However, as noted in the previous section, the 2008-09 evaluation the number of unit assessments completed did not correlate with higher achievement. From the 2008-09 evaluation report:

It does not appear that the number of unit tests completed or the unit test average provide meaningful information related to student progress toward their math goal. Neither performance indicator is related to growth on the MAP test.

This suggests that the additional minutes of tutoring may be more important in driving student achievement than the change in threshold in the unit assessments, but both play a role. If this program moves to replication in other schools, the data suggests it would be wise to include both the 90 minute-per-week target and the 85% threshold as program components.

The Master Coach, in an interview with evaluators, reported that both are important. However, she noted that the increase to 90 minutes of tutoring per week gave students more depth in each unit. Tutors did not spend the extra time in “re-drilling,” she observed, but covered more content with students. She believes this more indepth approach is a large part of the increased success rate of students in 2009-10.



Performance on MAP and MCA tests, comparing students who received 12 weeks or more of tutoring in 2008-09 with similar students in 2009-10. The Pearson Chi-Square Test suggests that the differences shown in both measures are statistically significant (MAP $p = .019$, MCA-II $p = .000$). Note that the 2008-09 data is approximate while the 2009-10 data is more accurate, since the data included number of weeks in the program.

FIDELITY

Research question: Are the integrity/fidelity checks used in 2009-10 effective? Is there evidence that they have improved fidelity over 2008-09? To what extent can improvements in student performance in 2009-10 be attributed to higher fidelity? Do students who receive tutoring with more fidelity do better than students receiving less?

Impact of fidelity checks in 2009-10

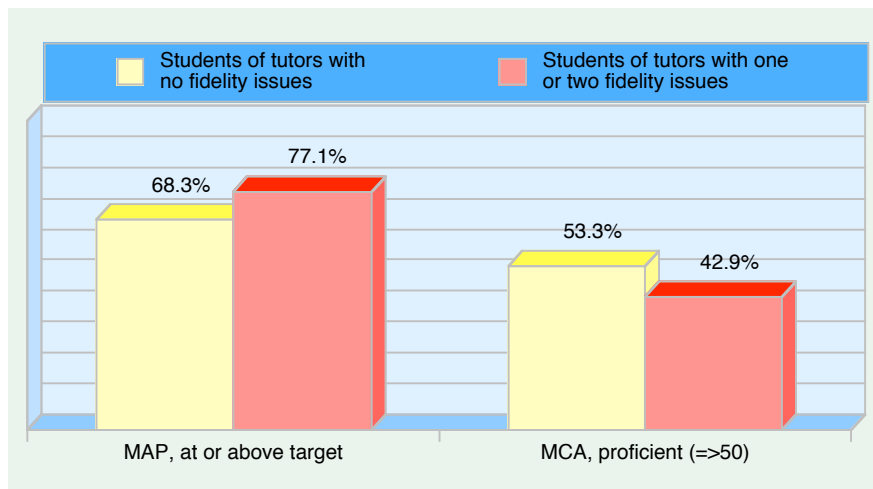
Math Corps used fidelity checks in both years. However, after feedback from an evaluator in 2008-09, the program made changes to increase the effectiveness of the process.

In 2009-10 both lesson observations and process checks were completed (each with their own observer guide/feedback form). Combining the process checks and lesson observations, an average of five fidelity checks were conducted with each tutor. One individual was checked ten times during the year.

In an interview with the researchers, the Master Coach described the more robust data collection process in 2009-10 resulting in quicker feedback to the member. She believes the process checks helped program staff to catch problems early, which led to a year with few issues. Overall, she reported that there were few examples of tutors not using the program in the way it was intended.

In examining the process and lesson observation forms (114 in all), there were only four instances (three individuals) where the “Need follow-up” box was checked.

An analysis of the outcomes for students of those tutors shows no significant difference between those with one or two fidelity issues and those without. The chart below shows the results, showing that students whose tutor had a few fidelity issues did not perform quite as well on the MCA testing, but actually did better on the MAP fall-to-spring RIT gain score.



Performance on MAP and MCA tests, comparing students whose tutor had no fidelity issues and students whose tutor had one or two fidelity issues. The Pearson Chi-Square Test suggests that the small differences shown are not statistically significant (MAP $p = .296$; MCA-II $p = .253$). There were 35 students tutored by the three individuals with one or two fidelity issues, and 195 students in the group of tutors with no issues.

The statistical tests, however, suggest that the small differences are not significant.

An additional test examined the mean RIT gain score in relation to each student's expected growth target, comparing the two groups. Those who were in the group with tutors who had one or two fidelity issues actually did significantly better than the larger group of all other tutors.

This supports the statement by the Master Coach that the few problems were caught early. The data suggests that the fidelity checks implemented in 2009-10 are effective.

The t-test showed that the mean score of 6.98 for the group with fidelity issues was significantly higher than the mean of 2.90 for the group with no fidelity issues ($p = .006$).

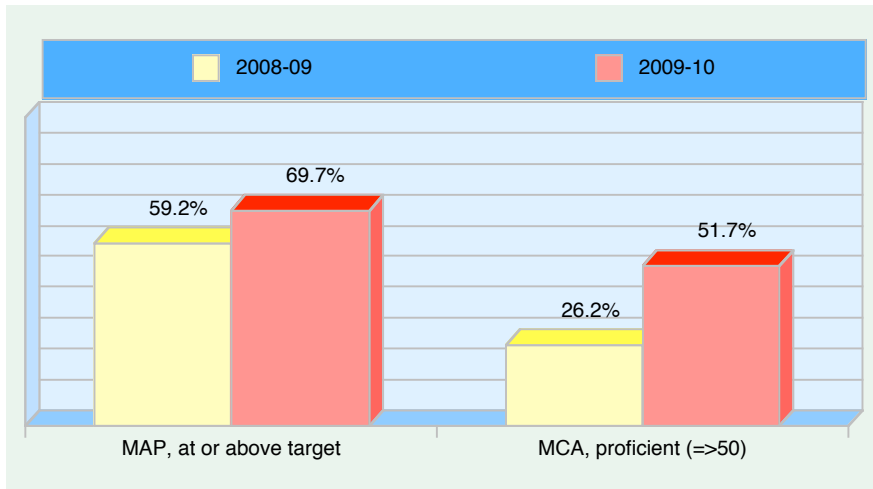
Comparison between 2008-09 and 2009-10

The process for measuring fidelity through observation was strengthened in 2009-10, because of suggestions from the evaluator in 2008-09. Direct comparison of fidelity is difficult between the two years. This is due to three factors:

- The records available for 2008-09 do not indicate which tutors were implementing the program with fidelity and which were not.
- In 2009-10, only three tutors had minor fidelity issues, a very small sample. This no doubt contributed to the data that showed no significant difference in student performance between students of tutors with fidelity issues and those whose tutors had no issues.
- In 2009-10, the goal for number of minutes each student was tutored was raised from 60 minutes per week to 90 (or higher for junior high), and the threshold for progression to the next unit was raised from 75% to 85%. The combination of more contact time and a higher threshold is likely to have produced gains in performance, according to the data described earlier. It is difficult to tell how much of the increase in performance is attributable to increased fidelity and how much to these other factors.

The chart on the next page show the overall results for the two years. The comparisons show the percent of participating students who were rated proficient or above on the MCAs and the percent who met or exceeded their growth target on the MAP assessments.

The data shows that students did better in 2009-10 than in 2008-09. This is likely the result of many factors, but increased fidelity in how the program was delivered may play a part. The higher success of the program in 2009-10, coupled with the revised fidelity process, suggests that individuals replicating this program in other locations plan a rigorous program to insure fidelity.



Performance on MAP and MCA tests, comparing students who participated in the program in 2008-09 and in 2009-10. The Pearson Chi-Square Test suggests that the differences shown in both MAP and MCA are statistically significant (MAP $p = .038$; MCA-II $p = .000$).

FACTORS THAT INFLUENCE ACHIEVEMENT

Research question—What factors significantly influence achievement? Do students of a specific ethnicity do better than others? Does socio-economic background significantly affect achievement? How is achievement different for students for whom English is a second language?

Ethnicity

The chart below shows the results for students of color compared to white students. All races except those identified as “white” were combined into the category “students of color.”

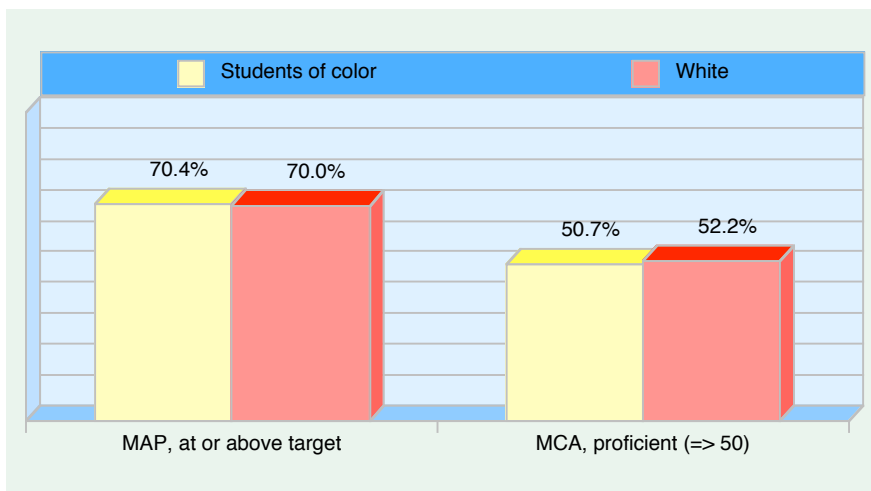
Both the MAP gain (at or above target) and the MCA-II (proficiency, score of 50 or above) are shown. Neither of the small differences are significant.

Black students are a majority of the students of color. Looking only at black students compared to white students, the data does show a difference in the percent proficient on the MCA-II test, but the statistics do not confirm the difference. And the percent of black students who met target in the MAP test between fall and spring was nearly identical to white students. The chart showing this data is on the next page.

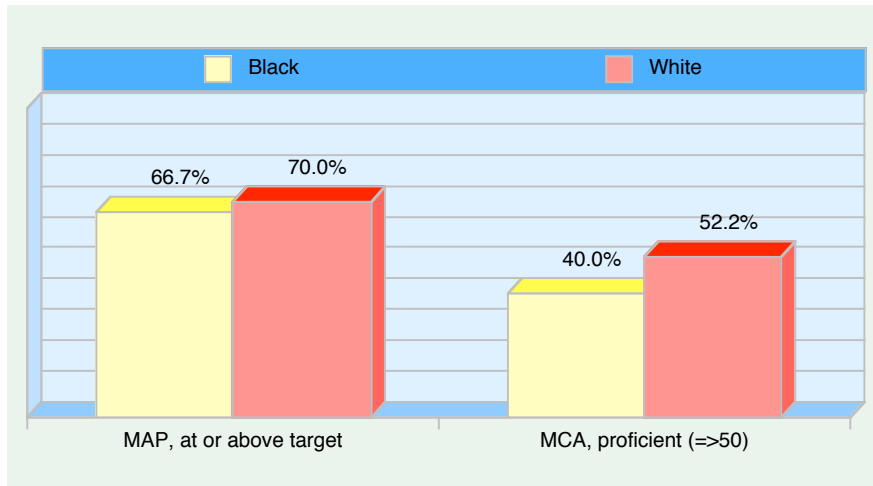
Note that the MAP gain score may be a slightly better measure of progress than the absolute proficiency scale of the MCA-II tests, since the MAP gains are normed for the different starting points of each student.

A t-test of the difference in the mean fall to spring gain score of students of color compared to white students also shows no significant difference in performance between the two groups.

Overall, the data indicates that ethnicity does not have much impact on how well a student does in the program. Or, put another way, the data indicates that the program is as effective with students of color as it is with Caucasian students.



Performance on MAP and MCA tests, comparing students of color with white students. The Pearson Chi-Square Test suggests that the small differences shown are not statistically significant (MAP $p = .948$, MCA-II $p = .834$).

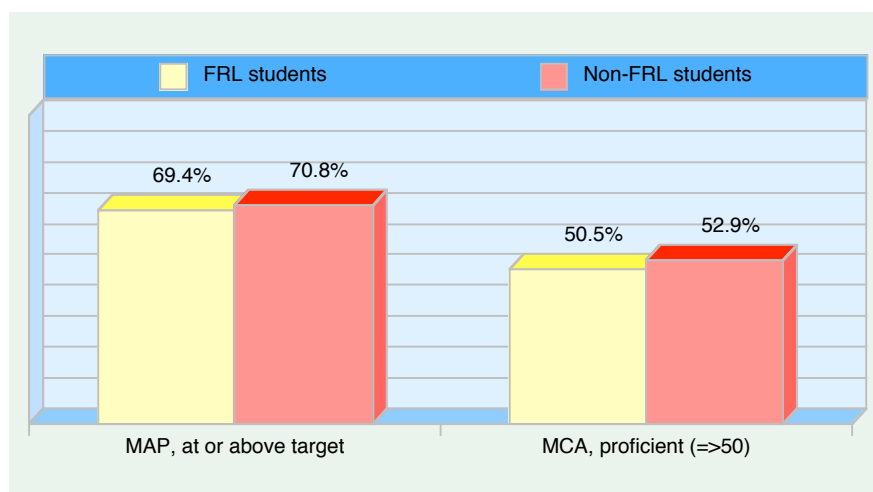


Performance on MAP and MCA tests, comparing Black students with White students. The Pearson Chi-Square Test suggests that the small differences shown are not statistically significant (MAP $p = .669$, MCA-II $p = .148$).

Socio-economic background

Students who are eligible for free or reduced lunch (FRL) come from lower income households. Families of students eligible for free lunches have income 130% or less of the poverty level (based on the Federal Poverty Guidelines); students eligible for reduced-price lunches are from families with incomes from 130% to 185% of the poverty level. 45% of Math Corps students in 2009-10 are FRL students.

The chart below shows Math Corps student performance in 2009-10 by FRL status. Neither of the small differences are statistically significant, suggesting that the program is effective with students from both sides of the socio-economic divide. A t-test of the difference in the mean gain score (fall to spring) of FRL students compared to non-FRL students also shows no significant difference in performance between the two groups.



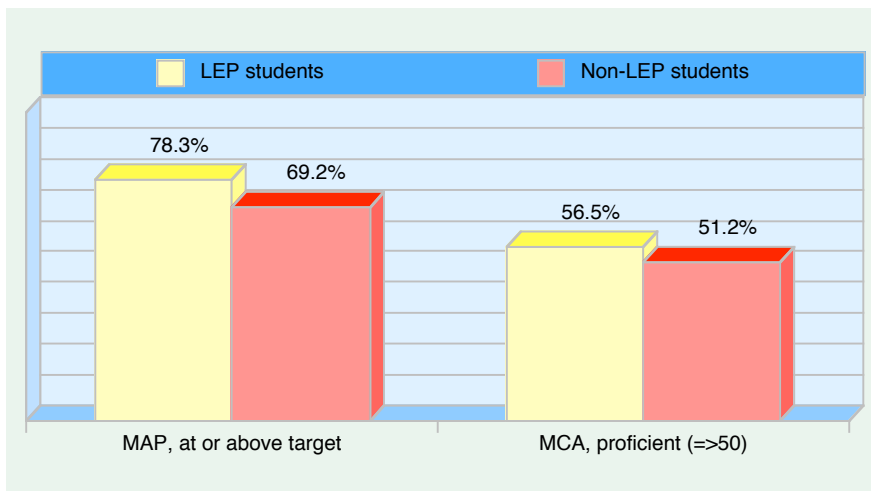
Performance on MAP and MCA tests, comparing FRL students (low socio-economic status) with non-FRL students. The Pearson Chi-Square Test suggests that the small differences shown are not statistically significant (MAP $p = .808$, MCA-II $p = .706$).

English Language Learners

About one in ten (9%) of Math Corps students in 2009-10 have Limited English Proficiency (LEP).

The chart below shows that this small group of LEP students performed slightly better than their peers who have higher English skills. However, the sample is small enough that the differences shown do not meet the test of statistical significance.

The data certainly suggest, however, that students who are English Language Learners benefit from Math Corps at least as much as those with better English skills. This is further confirmed by a t-test of the mean gain from fall to spring in MAP score; this test showed no significant difference between LEP students and non-LEP students.



Performance on MAP and MCA tests, comparing LEP students (Limited English Proficiency) with non-LEP students. The Pearson Chi-Square Test suggests that the small differences shown are not statistically significant (MAP $p = .369$, MCA-II $p = .629$).

Students of color	.315
White students	.365
t statistic	-0.434
Significance	.0665
FRL students	.343
Non-FRL students	.356
t statistic	-0.124
Significance	.0902
LEP students	.327
Non-LEP students	.352
t statistic	-0.142
Significance	.0888

This table shows the average amount that Math Corps students exceeded their expected gain score (the difference between the fall and spring scores) on the MAP test. Note that in all cases, the two groups of students had average gains well above their expected gains. Further, the t probability shows that the groups both experienced these gains at about the same rate.



IMPACT OF SECOND-YEAR TUTORS

Research question: Do students who work with second-year tutors do significantly better than students with first year tutors?

Second-year tutors

There were two second-year tutors out of the 22 total in the 2009-10 school year, in two separate schools.

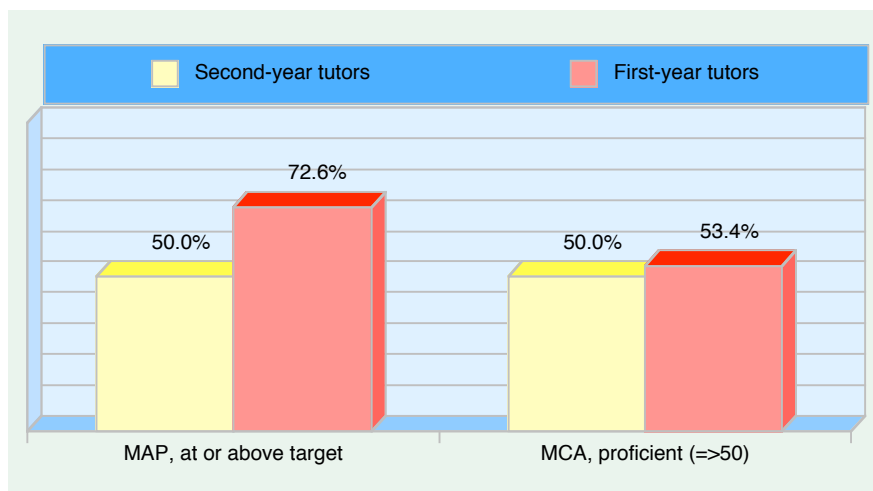
A question in an online survey of all tutors at the end of the school year asked those in schools with a second-year tutor to comment on the impact of having those individuals in the school. All five first-year tutors who had a second year tutor in the building rated it “Very helpful.” Asked to comment on how the program could make better use of second-year tutors, members suggested intentional pairing of first and second-year members, placing them in the larger schools (with the most members), and using them more during initial training.

In an interview, the master coach described both the second-year tutors in 2009-10 as “outstanding.” She believes that the self-confidence the two returning members showed was an important part of their success. The coach also noted that the initial training is “intense,” so the returning members have an easier time mastering the content the second time around.

Students of second-year tutors

Interestingly, the data shows that students of second-year tutors did not do as well as students of first-year tutors.

Note that the sample size is small; the two second-year tutors only worked with 20 students who had 12 or more weeks of tutoring. However, the statistics show that only 10 of these 20 students (50%) scored at or above their growth target in the spring MAP test, compared to about 73% of students of first-year tutors.



Students of second-year tutors were significantly less likely to meet or exceed their MAP growth target (the probability statistic was 0.35). The small difference in MCA proficiency was not significant ($p = .951$).

Analysis of the amount of gain from fall to spring compared to the norm also shows that students of second-year tutors did not do as well. The second year students on average scored 0.11 above their growth target while the rest of the students averaged 4.09. The difference is significant.

The probability for this t-test was .040.

It may be that the second-year tutors worked with more difficult students. Students of second-year tutors received only 66% of scheduled time, compared to 71% for the rest of the students. This would indicate more absences—either absences from school or schedule conflicts within the school—that impacted these students. In addition, two of the students who worked with one of the second-year tutors have missing data and had to be excluded from this analysis; this may also impact this finding.