EXAMINING MATHEMATICS COACHING

PRELIMINARY REPORT:
IMPACT OF PARTICIPATION IN EXAMINING MATHEMATICS COACHING (EMC) IN (K-8) CLASSROOMS ON STUDENT MATHEMATICS ACHIEVEMENT IN IDAHO

FUNDING BY THE NATIONAL SCIENCE FOUNDATION
DISCOVERY RESEARCH K-12 PROGRAM (DR K-12),
AWARD NO. 0918326

PRELIMINARY REPORT:
IMPACT OF PARTICIPATION IN EXAMINING MATHEMATICS COACHING (EMC) IN (K-8) CLASSROOMS ON STUDENT MATHEMATICS ACHIEVEMENT IN IDAHO

JOHN T. SUTTON, ELIZABETH A. BURROUGHS, DAVID YOPP,
DAN JESSE AND FELICIA FREEMAN

MARCH 1, 2011

Acknowledgments

This report is funded through a grant from the National Science Foundation,
Discovery Research K-12 Program (DR K-12), Award No. 0918326.
Any opinions, findings and conclusions or recommendations expressed in this material
are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Suggested citation: Sutton, J.T., Burroughs, E.A., Yopp, D., Jesse, D. and Freeman, F. (2011). *Preliminary Report: Impact of Participation in Examining Mathematics Coaching (EMC) in (K-8) Classrooms on Student Mathematics Achievement in Idaho.* Bozeman, MT: Montana State University and Denver, CO: RMC Research Corporation

For questions about this report, please contact the Examining Mathematics Coaching project
at (877) 572-5032 or email to: emc@math.montana.edu

Table of Contents

TABLE OF EXHIBITS

IMPACT OF PARTICIPATION IN EMC

BACKGROUND

The Examining Mathematics Coaching (EMC) project, funded by the National Science Foundation, conducts research on knowledge that contributes to successful coaching in two domains: Coaching Knowledge and Mathematics Content Knowledge. Additionally, the research examines impacts on teachers' knowledge, attitudes, and classroom practices based on their coaches' knowledge of coaching and mathematics content. The project also assesses impact on student achievement in selected settings. This partnership-driven project between Montana State University and RMC Research Corporation is designed to enhance the knowledge and understanding of mathematics coaching as they contribute to developing teacher quality, resulting in enhanced student learning in mathematics.

The EMC intervention consists of extensive coaching with K-8 teachers. Coaches conduct at least eight coaching sessions per teacher per school year; four of the sessions are in the areas of number sense and operations. Coaches also conduct sessions that follow the EMC coaching model (15 minute pre-observation conference, observation, 30 minute post-observation conference).

The intent of this project is to advance the understanding of coaching effectiveness, while promoting teaching, training, and learning through coaching as a professional development model. This brief report presents preliminary results related to the impact on students in classrooms taught by teachers who have participated in the project, compared to the performance of students taught by other teachers in the same schools.

METHODOLOGY

One of the activities of the EMC coaching project is to look at any possible impacts on student achievement of teachers who are being coached through the project. A pool of 46 teachers in the state of Idaho participated in the project. Efforts were made to compare student achievement results from a subset of 16 of these 46 teachers with a comparison group of teachers generated through the student achievement database from the state of Idaho. These teachers had students who completed both the 2009 and 2010 Idaho Standards Achievement Test (ISAT) mathematics assessment. Results included students who were in the third through sixth grades in 2010. The EMC project began in July 2009, so the 2008-2009 student achievement data were used to establish a baseline, and the 2009-2010 student achievement data were used as a first year comparison. Only teachers who were enrolled in the project before the test was administered in the spring of 2010 were included in this preliminary analysis. This report describes preliminary findings from analysis of ISAT mathematics results, and puts forth suggestions for further in-depth exploration of the data to inform project activities.

Students in Grades 3-8 and Grade 10 take the ISAT each spring. Due to the population of students in participating classes, sample characteristics, and the fact that scores were needed from 2009 and 2010 for each student, seventh and tenth graders were not included in these analyses. For ease of interpretation and to allow for pooling across grades, ISAT scale scores were converted to normal curve equivalent (NCE) scores.

Participating teachers who had students with matched data sets were included in these analyses. The sample drew from five districts, 13 schools, and 264 students of 16 teachers who worked with nine coaches. The comparison sample of the rest of the students in the five districts consisted of 16,663 fourth, fifth, sixth and eighth graders. In order to identify students for this comparison group, student data were used from the same five districts as the teachers who were trained in the program. Comparison of demographic information revealed that this approach was successful in identifying classrooms for the contrasts planned. Exhibit 1 displays demographic characteristics of students in classes taught by participating teachers, and comparison students in the same districts taught by nonparticipating teachers. This Exhibit reveals that except for size, the convenience comparison sample is very similar to the intervention sample. There are somewhat higher percentages of students on free or reduced price meals in classrooms of the participating teachers sample than in the other schools in the same districts, otherwise the characteristics of the students from the two groups are basically the same.

EXHIBIT 1. CHARACTERISTICS OF STUDENTS OF PARTICIPATING AND COMPARISON GROUP TEACHERS

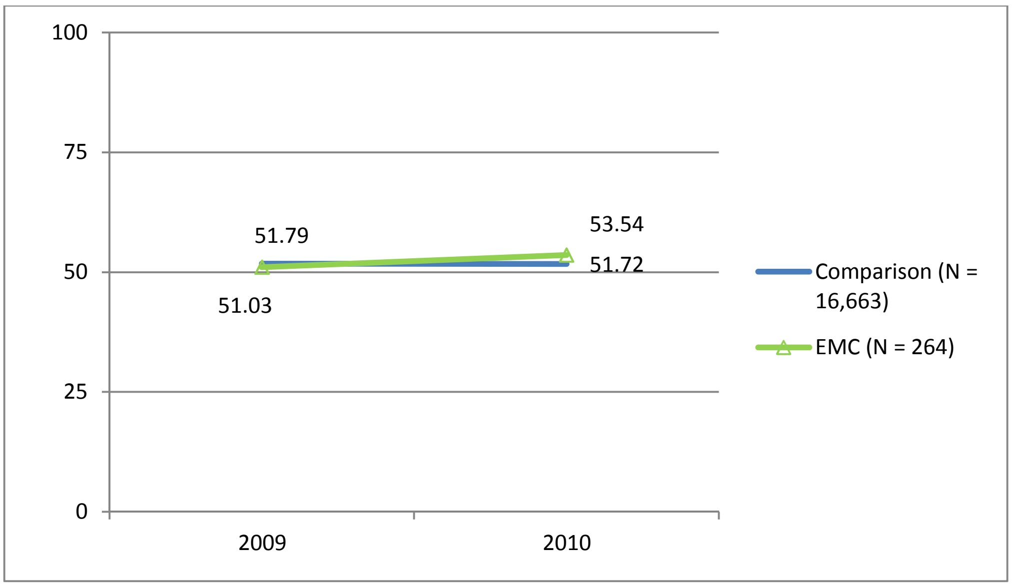
Characteristic	Percentage of Students of Participating Teachers (*N* = 264)	Percentage of Students of Comparison Group Teachers (*N* = 16,663)
Male	51.5	51.3
Female	48.5	48.7
Native American	4.2	1.9
Asian	2.3	1.8
African American	1.1	1.2
Pacific Islander	0.8	0.3
White	82.6	83.4
Hispanic	8.0	9.5
More than 2 races	1.1	1.9
Free/Reduced Price Lunch	47.3	36.5
Limited English Proficient	2.3	3.5
Special Education	11.4	10.2

FINDINGS

The primary question of interest was whether teacher participation in the project had an impact upon the mathematics achievement of their students. This question was answered generally, and then more specific findings were explored in some detail. Exhibit 2 displays the results for a 2 (participating teachers vs. comparison teachers) by 2 (2009 ISAT mathematics vs.

2010 ISAT mathematics) repeated-measures ANOVA.[1] Students of teachers in the project increased their mathematics achievement scores in 2010, while comparison students did not, $F(1, 16) = 9.833$, $p = .002$.[2] However, the effect size for this result is very small.[3] Results are displayed graphically in Exhibit 2 and reveal only a slight advantage for students of participating teachers.

Exhibit 2. Student ISAT Mathematics Scores for 2009 and 2010 by Teacher Participation



While the results look relatively straightforward, they are much more complex than the analysis revealed. A sizable proportion of variance is not accounted for by this simplified analysis which does not take the hierarchical nature of the data into account. To more adequately test the differences between the groups of students, a hierarchical ANOVA[4] design with teacher participation nested within districts was used. Results indicated that this nesting relationship across time was statistically significant, $F(5, 16{,}917) = 5.550$, $p < .001$, thus indicating that there was considerable variability between school districts as might be expected. Exploratory analyses also revealed wide variation between participating teachers as a function of their schools, and as a function of their coaches. Exhibit 3 displays variation between schools, and

[1] Repeated measures analysis of variance (ANOVA) is a statistical procedure used to examine differences within subjects when the same data are available for multiple points in time.

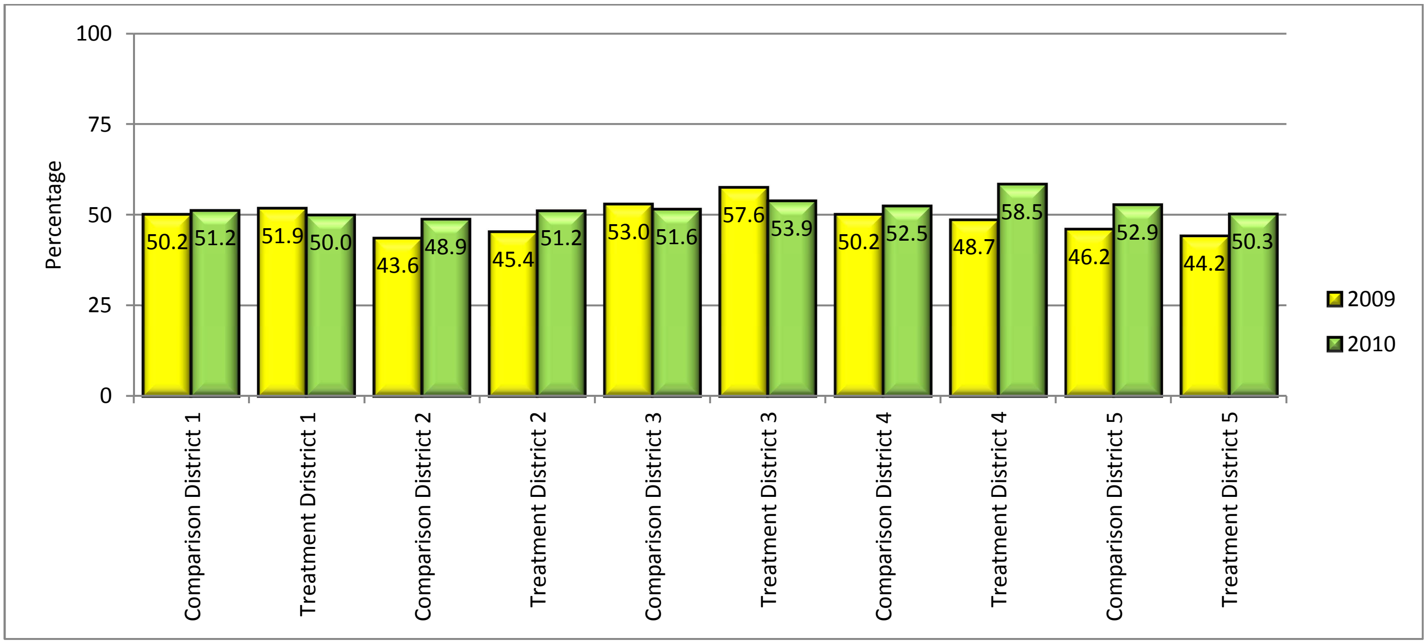
[2] The *F* statistic provides a basis to test for statistical significance when used in analysis of variance (ANOVA). The *p*-value is an indicator that represents the likelihood that observed results occurred by chance. In education research, values of $p < .05$ (i.e., values indicating that observed results had a less than 5% chance of occurring by chance) are typically used to identify results that are statistically significant. Lower *p*-values indicate a smaller likelihood that observed results occurred by chance and are therefore associated with statistically significant findings.

[3] Generalized Eta Squared for this analysis is less than .001. Eta squared (η^2) is an effect size that measures the proportion of variance of a dependent measure that can be explained by an independent variable. Eta squared statistics are commonly used in analysis of variance (ANOVA) and *t*-test analysis and their values range from 0 to 1.

[4] A nested ANOVA or a hierarchical ANOVA accounts for variability unique to subjects in groups. For example, students are usually nested within classrooms, which is a hierarchical data structure that should be adjusted for statistically.

indicates that results were the most promising in District 4. Students of teachers who participated in EMC in District 4 made larger mathematics achievement gains in the ISAT than did the other students in the same district.

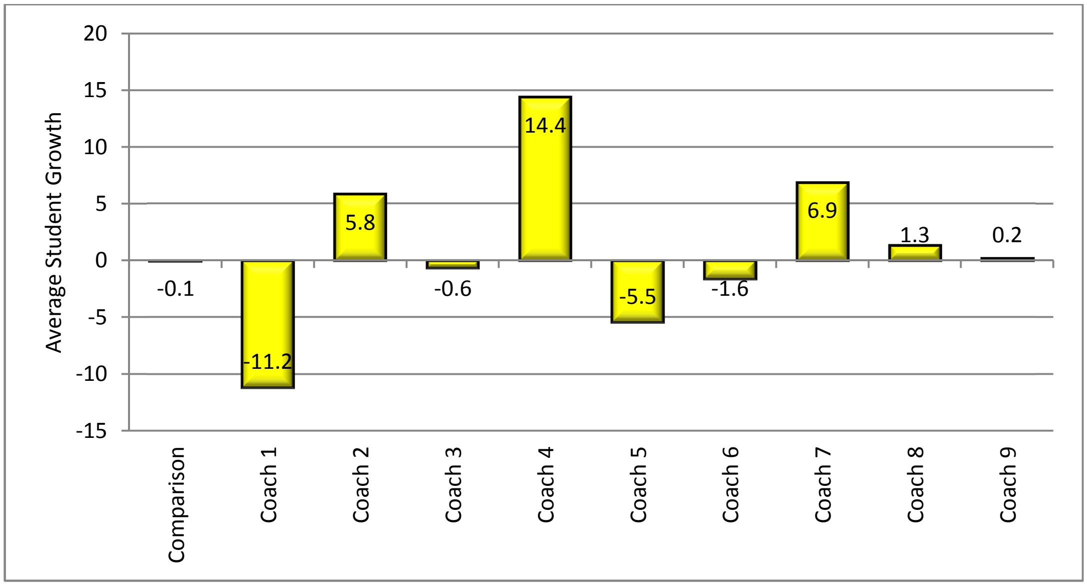
Exhibit 3. Student Achievement Results by District Disaggregated by Teacher Participation



Note. Treatment and comparison groups are nested within five school districts.

A total of nine coaches worked with the 16 teachers in the project who were studied here. For ease of interpretation, difference or growth scores were calculated by subtracting 2009 scores from 2010 scores for each student in the sample. Exhibit 4 displays these growth scores by coach, and indicates the nature of the variability between coaches. Coach 4, Coach 7 and Coach 2 worked with teachers whose students grew more in achievement than did the comparison group or other coaches. These results should be interpreted with some caution, because coaches worked with no more than two teachers in the sample.

EXHIBIT 4. STUDENT ACHIEVEMENT RESULTS BY DISTRICT DISAGGREGATED BY TEACHER PARTICIPATION



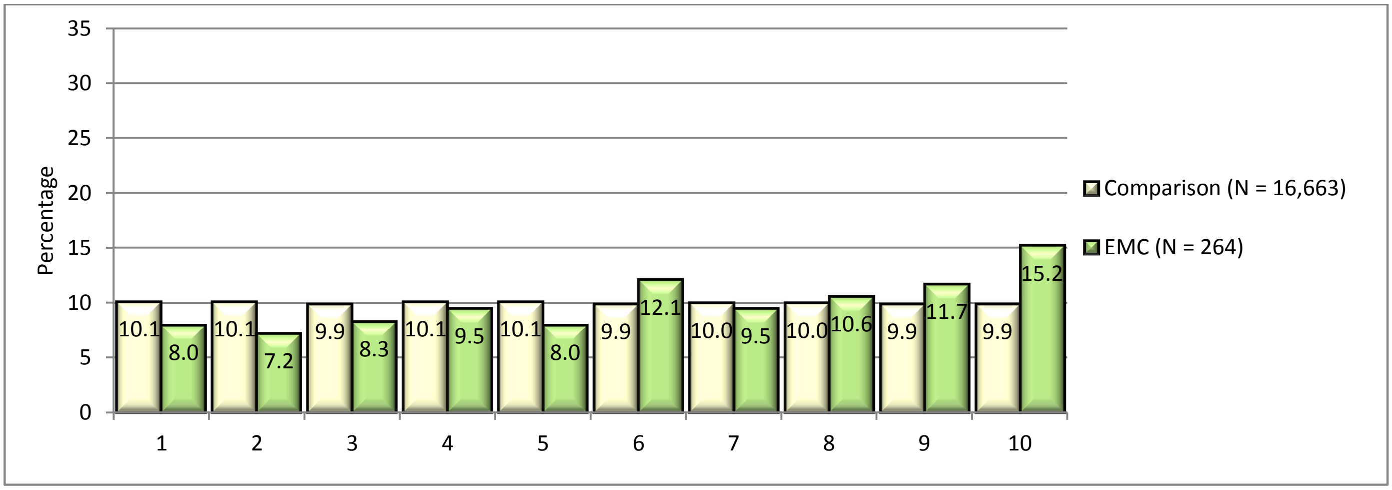
BINNING ANALYSIS OF MATHEMATICS STUDENT ACHIEVEMENT RESULTS

As a proxy for a value-added model which quantifies the impact teachers have upon student achievement, growth scores from 2009 to 2010 were calculated. Next, these growth score values were "binned" by placing them into 10 equal categories. Generally speaking, lower bin numbers were indicative of less growth, while higher bin numbers revealed higher levels of growth from one grade level to the next.

Students of teachers who were in the project were compared to students of other teachers in the same school districts. These comparisons were made for students in the fourth, fifth, sixth, and eighth grades who took the ISAT mathematics assessment in 2009 and in 2010, and whose records could be matched.

Exhibit 5 displays the results of this binning procedure for the students of participating teachers compared to other students in the same districts. This Exhibit suggests that students of teachers working with coaches participating in EMC grew more on the ISAT from 2009 to 2010 than did other similar students during the same time period. This was particularly true for the students at the highest bin levels.

EXHIBIT 5. BINNED RESULTS FOR PARTICIPANT TEACHER STUDENTS AND COMPARISON TEACHER STUDENTS



Next, binning results were broken down by selected grade levels reflecting the data available. Exhibit 6 reveals that there was some variation between the two groups with regard to binning. The expected value of a bin is about 10% of the students in a group. Again, bins are ordered in terms of student growth from third grade to fourth grade in this example. Bins on the left indicate negative or low growth, while bins on the right indicate positive or high growth. Exhibit 6 revels that students of teachers working with coaches in the EMC project scored higher than other students in the same school districts in bins 9 and 10, where the highest gains were made from third grade to fourth grade.

EXHIBIT 6. BINNED FOURTH GRADE RESULTS FOR PARTICIPANT TEACHER STUDENTS AND COMPARISON TEACHER STUDENTS

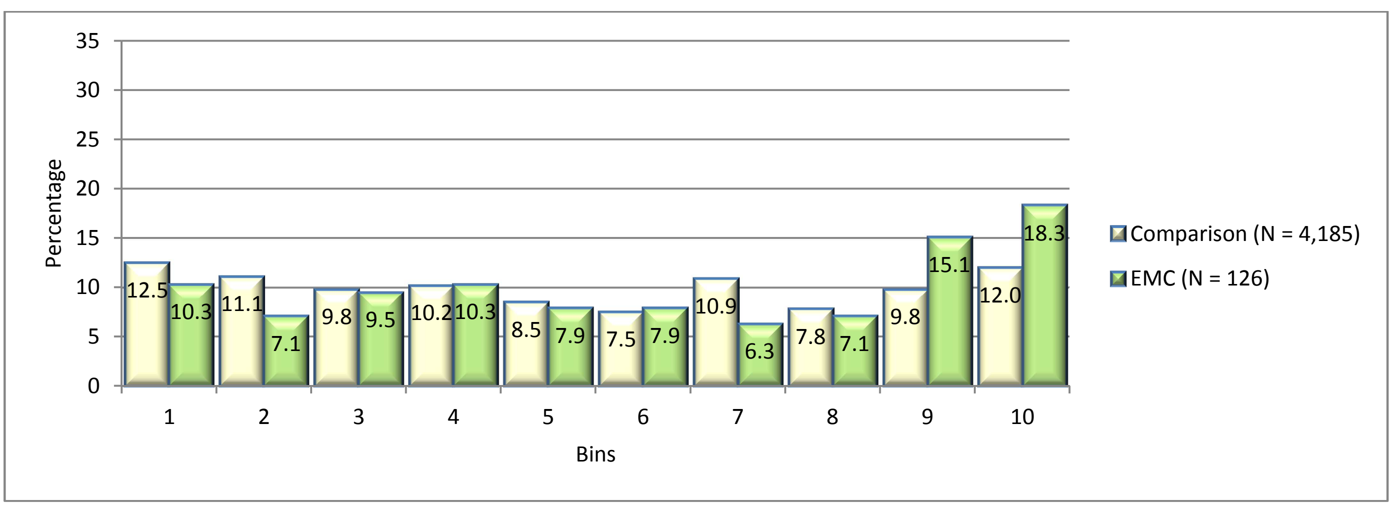
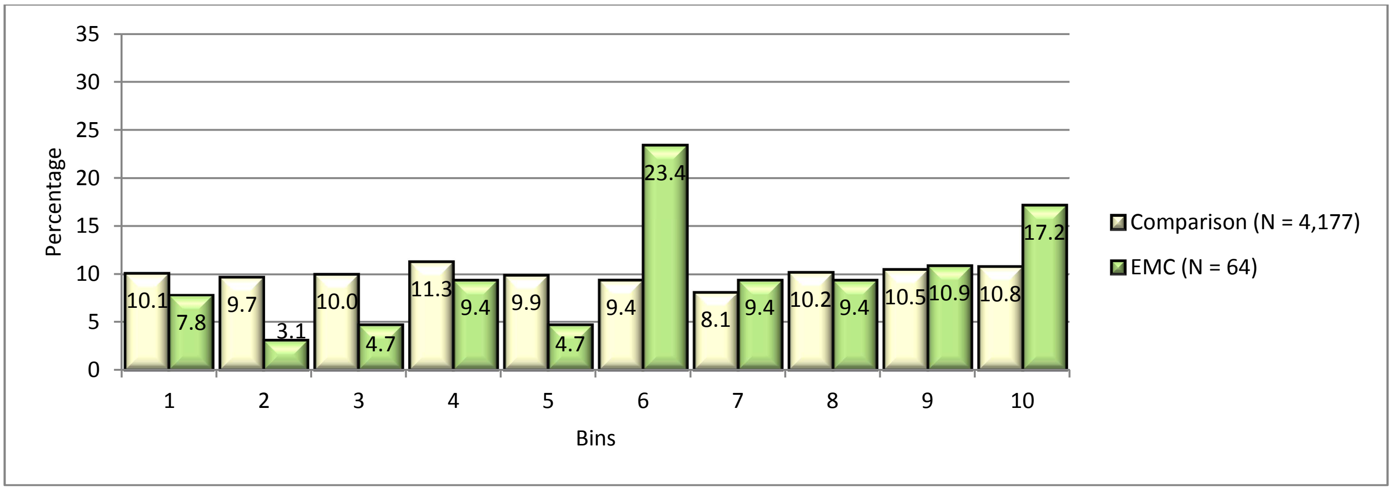


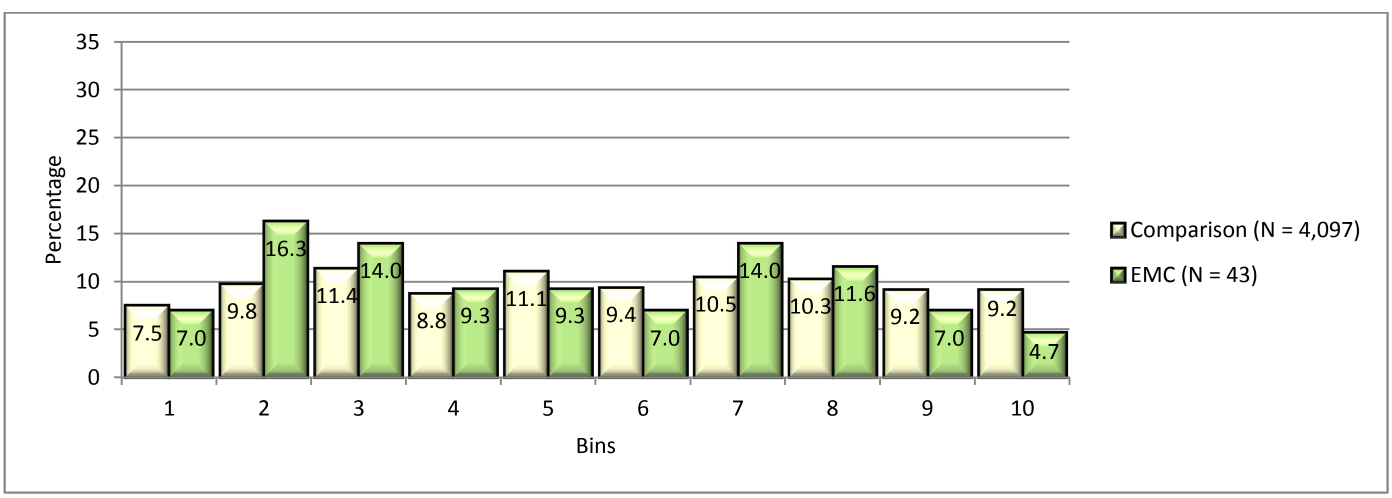
Exhibit 7 displays binned results for fifth graders. While results are not conclusive, there is the suggestion that students in the highest growth bin were taught by teachers working with coaches in the EMC project in 2009-2010, and that smaller percentages of students of teachers working with coaches in the EMC project were in the lower performance bin categories.

EXHIBIT 7. BINNED FIFTH GRADE RESULTS FOR PARTICIPANT TEACHER STUDENTS AND COMPARISON TEACHER STUDENTS



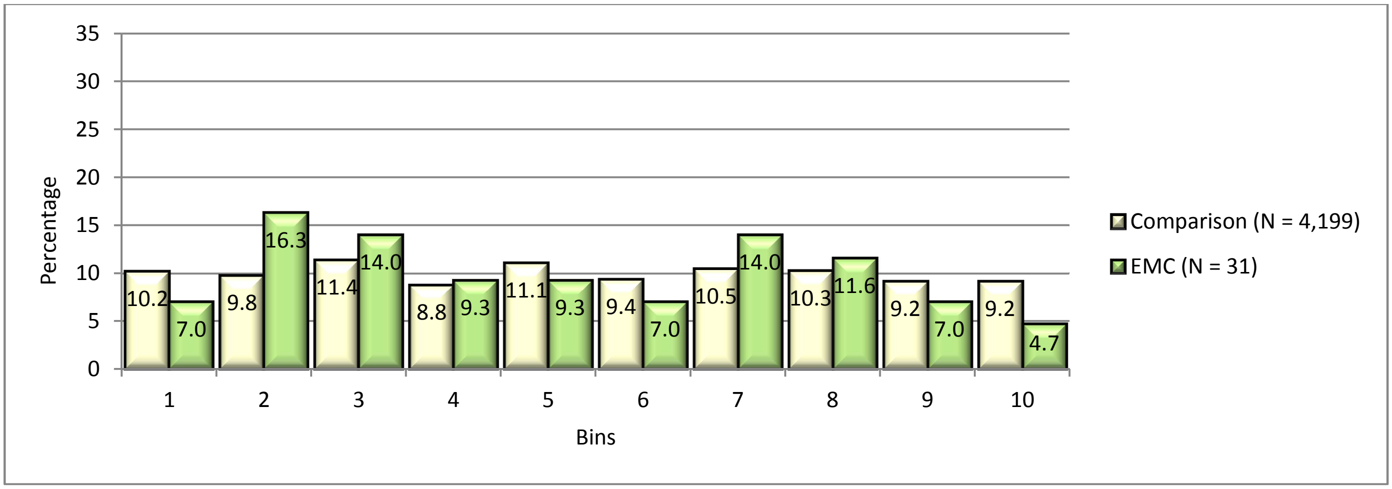
Sixth-grade students of teachers working with coaches in the EMC project did not seem to perform much differently than comparison group students as evidenced by Exhibit 8, which suggests higher numbers of the students of participants in some of the lower bins, and lower numbers in the highest bins. Results should be interpreted with some caution because there were only 43 students of participating teachers in this analysis.

EXHIBIT 8. BINNED SIXTH GRADE RESULTS FOR PARTICIPANT TEACHER STUDENTS AND COMPARISON TEACHER STUDENTS



Eighth-grade students' results in Exhibit 9 show no consistent pattern of results favoring students of participating teachers or students of other teachers in the same districts. Again, these results should be interpreted with some caution because of small sample sizes of students of teachers working with coaches in the EMC project.

EXHIBIT 9. BINNED EIGHTH GRADE RESULTS FOR PARTICIPANT TEACHER STUDENTS AND COMPARISON TEACHER STUDENTS



IMPLICATIONS

The results of this preliminary study are promising. Overall, students of teachers who worked with coaches participating in the EMC project increased their test scores slightly more than students of other teachers who did not work with coaches. However, the differences are slight, and there is considerable variation in the data that are not explained by just looking at students of participating teachers. Exploratory analyses revealed that there is considerable variation between districts, between coaches, and between classrooms, which should be accounted for in further analyses using hierarchical linear modeling (HLM).

Further investigation is warranted to determine the nature and causes of the variation between districts and coaches. For example, it is important to know why project activities were associated with success in District 4. It is also important to understand more about why student achievement growth was so pronounced for the teacher(s) who worked with Coach 4.

Finally, teacher variation should be systematically explored. First, teachers ranged in experience with the project. Some signed up in October of 2009, while others did not enroll until November or later. The project routinely collects teacher knowledge information and demographics about teachers, which could be used for such explorations.