Relating Instructional Practices in Mathematics to Student Success: Focus on Math 7 for Grade 7

Rachel A. Hickson, Evaluation Specialist, Montgomery County Public Schools
Shahpar Modarresi, Supervisor, Program Evaluation, Montgomery County Public Schools

Abstract

In a large Maryland public school district where all middle school students are expected to complete Algebra 1 before high school, this study sought to identify instructional practices in Math 7 classes that have positive and significant associations with students' mathematics achievement.

A multi-method evaluation design included a formative component to collect and analyze multiple classroom observations of instructional practices, and a summative component which used advanced statistical models to examine relationships between observed instructional practices and student success in mathematics. Several practices were found to have a significant positive relationship to student mathematics performance. Results will inform mathematics teaching at many levels and to different types of students.

Introduction

The curriculum office in Montgomery County Public Schools (MCPS), a large and highly diverse Maryland school district near Washington, DC, requested a study to examine whether teacher use of recommended instructional practices has a significant positive relationship to academic performance. The focus course for this study was Math 7 in Grade 7, a course for non- advanced mathematics students in United States Grade 7 (students are 12 and 13 years old). The instructional practices examined in this study are believed by experts to support successful learning regardless of the mathematics course level. Results are intended to inform the teaching of mathematics at a variety of levels and to a broad population of students.

Method

This study utilized a multi-method data collection strategy, conducting both formative and summative studies. A formative study was conducted by collecting and analyzing observations of Math 7 classes. A summative study relied on a quasi-experimental design (Shadish, Cook & Campbell, 2002) to maximize internal validity by controlling confounding variables. Advanced statistical techniques were used to address confounding variables and improve internal validity.

Formative Study

The formative study was designed to answer the following question: *Are Math 7 teachers of Grade 7 students using recommended instructional practices?* The formative study was conducted by collecting and analyzing observations of Math 7 for Grade 7 classes.

All teachers of the Math 7 for Grade 7 course for at least two of the past three years (2010–2011, plus 2009–2010 and/or 2008–2009) were selected for observation. The final sample included 45 teachers working in 32 middle schools.

An observation protocol was designed in consultation with school district mathematics experts. Resources from the educational literature as well as district references were considered in designing observational indicators (MCPS, 2003; MCPS, 2007). Some observation indicators were repeated from past mathematics evaluations conducted in the district, including Algebra 2 implementation (Hickson, 2010).

Indicators were drawn from practices of interest, including emphasis on critical thinking and questioning; use of multiple strategies, materials, and modalities; use of classroom technology; use of small groups; use of discourse (teacher-student, student-student); evidence of classroom structures that support learning; and use of formative assessment and checking for understanding.

Additional steps were taken to ensure a relevant, high-quality instrument. The protocol was pre-tested in a non-sampled class, updated based on pretest experience, then reviewed again by mathematics specialists.

Two observations with each selected teacher took place during Unit 2 of the Math 7 course (fall 2011), with the first observation in late October/early November ("Time One"), and a second period about four weeks later ("Time Two"). Brief data collection windows help control for changes in the learning environment over time.

Mathematics class length varies by school. The average length of all observed classes was 58 minutes, with a range of 42 minutes to 95 minutes. Observers observed for the entire class period, or for 50 minutes, whichever was shorter. To ensure comparable observations at Time One and Time Two, observers of classes lasting longer than 50 minutes observed the first 50 minutes of the period at Time One and the last 50 minutes of the period at Time Two.

To account for additional instances of use of instructional practices on non-observation days, teachers completed a description of the four lessons or lesson sequences leading up to the observed lesson. These "lesson logs" prepared by observed teachers were later analyzed to determine whether recommended instructional practices were among those described in the logs.

School district instructional specialists reviewed lesson topics and handouts used in observed classes to determine whether content and topics were within the scope of the Math 7 course. Descriptions of the topics of observed lessons were classified into categories. Since self-reporting can inflate reporting of recommended practices, observation data helped to validate the self-reported information.

Summative Study

The methodology for two questions addressed in the summative study is summarized below. *Are recommended instructional practices used by Math 7 teachers significantly related to student outcomes, as measured by Maryland School Assessment (MSA) mathematics?* For measuring the use of recommended practices, an observation instrument identifying specific indicators of instructional practice was developed and used in Math 7 classrooms (see formative study, above). The recommended practices in the observation instrument measured two parts. In part one, several indicators in the instrument were grouped into four categories or constructs, measuring (frequency of observation) the following: a) critical thinking and questioning; b) differentiation, variety, and learning styles; c) classroom technology; and d) formative assessment. In part two, several dichotomous indicators were used to measure the presence of the following practices: a) discourse and group work, b) classroom structure, and c) exit card/summarizer. The outcome measure (dependent variable) for addressing this question was student performance on Grade 7 MSA mathematics. MSA is a standardized test that demonstrates how well Maryland students have learned the skills specified in the state curriculum. The test is administered annually in Grades 3 through 8 and serves as Maryland's accountability measure under federal law. In this study, Grade 7 MSA scores were used as the outcome measure and Grade 6 MSA scores were used to control for students' initial abilities in statistical models. The correlation between Grade 7 MSA and 6 Grade MSA was significant ($r = .84$; $p < .001$). MSA mathematics scores were used to address both summative questions.

Several statistical analyses were performed on groups or constructs of practice (based on the number of times each practice was observed). First, exploratory factor analysis (principal component) was applied to indicators in each group of the instructional practices (or index) in Math 7 classrooms. Factor scores were then placed in variables and saved in the data set for the purpose of multiple regression analysis. Second, the coefficient alpha was computed separately for each factor to ascertain the reliability of the measures (the extent to which measures making up each factor share a common core). Third, multiple regression procedures were used to examine whether the better Grade 7 students' mathematics outcome would be significantly associated with the higher use of the recommended instructional strategies or indicators of the Math 7 classroom practices. For the dichotomous observation indicators (coded as 0 and 1), only multiple regression analyses were used to test these indicators' associations (negative or positive) with students' MSA mathematics scores. Multiple regression analytical procedures were performed separately for each factor (or group) of observation indicators as well as individual indicators of the practice.

The internal validity of the findings was improved by controlling for students' prior performance on mathematics, teachers' years of experience, students' initial abilities, demographics, and service receipt measures in the statistical models. To avoid the problem of multicollinearity, the factor analytical procedures (principal component) were applied to student background information to create two orthogonal factors. Background information included students' initial abilities, demographics, and service receipt measures. The teachers' experience measure was removed in the final analyses, since it did not explain significant variation in students' mathematics performance as measured by MSA.

Are there differences in mathematics performance between students of observed teachers (more recent experience teaching Math 7 for Grade 7) and students of non-observed teachers (less recent experience)? For this question, the outcome measure (dependent variable) was again the Grade 7 MSA mathematics scores. Analysis of Covariance (ANCOVA) was used to statistically control for the effects of possible pre-existing differences between the two groups of students (students of observed teachers vs. students of non-observed teachers). Effect size measures were used to examine the magnitude of mathematics achievement differences between the two groups of Grade 7 students.

The confounding variables were controlled for through the use of propensity scores as well as advanced analytical procedures. Propensity scores based on students' background characteristics (e.g., race, gender, and receipt of Free and Reduced-price Meals System [FARMS] or special education services) were computed using logistic regression models **(**Luellen, Shadish, & Clark, 2005)**.** The propensity scores were divided into five categories and used as covariates in the statistical models, which included prior ability (Rosenbaum & Rubin, 1983, 1984, 1985). Both statistical significance tests and effect size statistics were used to address the study question.

To address the question of relationship of practices to student outcomes, this study used indicators of recommended Grade 7 mathematics practices as captured by an observation instrument. The measures in the instrument were available for observation in every class. Four categories of indicators (groups of practices) were relevant to the second question, including: a) critical thinking and questioning; b) differentiation, variety, and learning styles; c) classroom technology; and d) formative assessment. The response format for the above indicators was a scale for number of observations, ranging from not observed during the lesson (0) to observed six or more times (6). Other indicators of Math 7 instructional practices included: a) three measures of Discourse and Group Work; b) four measures of Classroom Structure; and c) two additional indicators: "asking questions at a variety of levels (recall, comprehension, inference)," and "Exit Card/Summarizer." These indicators were operationalized as "0" (not observed in the lesson) or "1" (observed one or more times during the lesson). Each observed teacher and his or her students received a series of scores based on the observed use of the recommended indicators. Since each teacher was observed teaching twice (two different lessons on different days), scores were averaged for the two observations for each observed teacher and used for advanced statistical analyses.

Samples for the summative study. The analytical sample for the relationship of practice to student performance included students of 44 observed Math 7 teachers. The final sample included those students of observed teachers who had valid scores on both Grade 7 MSA and Grade 6 MSA mathematics. Students who took an alternative assessment were not included.

The analytical sample for the relationship of experience to student performance included students of both observed teachers (44) and non-observed teachers (43) who taught Math 7 for Grade 7 in FY 2011 . As a group, non-observed teachers had fewer years of recent experience teaching Math 7 for Grade 7; on average non-observed teachers had fewer years of overall experience.

Analytical procedures. The following analytical procedures for applied to address the formative question. Each observed instructional practice included in the Math 7 observation protocol was used for descriptive analysis (to add the number of times observed, to determine the percentage of classes in which particular indicators was seen, and so forth). Similar treatment was given to observations of lesson components and other lesson profile descriptors.

The following analytical procedures were applied to address the first summative question. Several statistical analyses were performed on groups or constructs of practice (based on the number of times each practice was observed). First, exploratory factor analysis (principal component) was applied to indicators in each group of the instructional practices (or index) in Math 7 classrooms. The factor scores were then placed in variables and saved in the data set for the purpose of multiple regression analysis. Second, the coefficient alpha was computed separately for each factor to ascertain the reliability of the measures (the extent to which measures making up each factor share a common core). Third, multiple regression procedures were used to examine whether the better Grade 7 students' mathematics outcome would be associated with the higher use of the recommended instructional strategies or indicators of the Math 7 classroom practices. For the dichotomous observation indicators (coded as 0 and 1) only multiple regression analyses were used to test these indicators' associations (negative or positive) with students' MSA mathematics scores. The multiple regression analytical procedures were performed separately for each factor (or group) of observation indicators as well as individual indicators of the practice.

To address the second summative question, Analysis of Covariance (ANCOVA) was used to statistically control for the effects of possible pre-existing differences between the two groups of students (students of observed teachers vs. students of non-observed teachers). Effect size measures were used to examine the magnitude of mathematics achievement differences between the two groups of Grade 7 students.

All multivariate analyses were guided by What Works Clearinghouse *Procedures and Standards Handbook* (U.S. Department of Education, 2008).

Findings

Formative Study

Are Math 7 teachers of Grade 7 students using recommended instructional practices? This section details findings from observations that provided critical data for exploring relationships between instructional practices and student performance. Table 1 summarizes the findings from observations.

The MCPS mathematics program provides curriculum look-fors to help teachers know what instructional practices are desirable and expected in their classes (MCPS, 2007). Teachers are expected to utilize them on an ongoing and regular basis. These practices should be apparent regardless of which class and which lesson component is observed, and by which observer(s).

To assess whether key instructional practices were being implemented, OSA evaluators observed 45 Math 7 teachers at 32 middle schools during Unit 2 of instruction. Each teacher was observed teaching the same class period twice: once in late October or early November (Time One, 45 classes) and again in late November or early December (Time Two, 44 classes). Data for one teacher not available at Time Two were dropped where appropriate.

To ensure that each class was observed for about the same amount of time, observers of classes lasting longer than 50 minutes observed the first 50 minutes of the period at the first observation and the last 50 minutes of the period at the second observation. This is particularly pertinent to observation of lesson components since warm-up is expected to occur only near the beginning of class and closure only near the end.

Classroom structures that support learning. Each observer looked for several indicators showing that classes had some established structures for supporting middle school learning. These indicators refer to the types of rules and routines that teachers put in place that allow students to be more independent and to reduce non-instructional time.

In a large majority of observed classes, students appeared to know what to do upon entering class, such as finding their seat, picking up work for the day, or finding their group members when they moved into a group (88.9% of classes at Time One, 88.6% at Time Two). In two thirds of all observed classes, ground rules or expectations were posted.

Remaining indicators of classroom structures that support learning were seen in less than one half of observed classes. Evidence that students dropped off completed work or picked up homework or make-up work without the teacher's help was apparent in 42.2% of classes at Time One and 36.4% of classes at Time Two.

Critical thinking and questioning. Observers looked to see whether teachers were using one or more indicators of critical thinking and questioning in their classes. Most teachers asked students questions that focus on problem-solving strategies and reasoning (97.8% of teachers at Time One; 90.9% at Time Two). A large proportion also modeled the thinking process for developing strategies and discovering relationship (95.6% of teachers at Time One; 84.1% at Time Two).

A sizeable majority of teachers reinforced students' use of the language of mathematics (84.4% at Time One; 72.7% at Time Two). About 7 out of 10 helped students make connections to prior knowledge (71.1% at Time One; 65.9% at Time Two). About two thirds of teachers presented or demonstrated multiple strategies to students (64.4% at Time One; 65.9% at Time Two). Finally, about one half of teachers used "real world" applications of mathematical concepts (51.1% at Time One; 47.7% at Time Two).

Discourse and group work. Observers looked to see whether and how teachers had students engage in discussions of mathematics, and whether they had students work or discuss in groups or pairs. Small group or partner work was observed in about one half of observed classes (53.3% at Time One; 45.5% at Time Two). Other indicators of discourse and group work were less common. Teachers facilitated student discussions about mathematical concepts and processes in a minority of classes (17.8% at Time One; 11.4% at Time Two). Also in a minority of classes, teachers had students discuss in groups or pairs (17.8% at Time One; 20.5% at Time Two).

Differentiation, variety, and learning styles. Several indicators measured teachers differentiating instruction for students and addressing multiple learning styles. In nearly all classes at Time One (93.3%) and a large majority of classes at Time Two (79.5%), teachers used a variety of materials and modalities to teach the lesson to the whole class (for example manipulatives, paper-and-pencil activities, technology, and discussion).

Just under one half of teachers encouraged students to try a variety of materials and methods to solve problems or generate responses (44.4% at Time One; 47.7% at Time Two). In about one third of classes, teachers varied activities, formats, or outcomes to support individual students' learning. (Some of this was accomplished by having a special educator assigned to the class.)

Formative assessment. Observers looked for a number of indicators that teachers were checking for understanding and using methods of formative assessment. (Observations were not scheduled on days when structured assessments took place.) Nearly all teachers (95.6% of classes at Time One; 100% at Time Two) asked direct questions to check for understanding and listen to students' responses. Nearly as many (93.3% at Time One; 90.9% at Time Two) did visual checks of students' work or homework at their desks (walking around to look at their answers, not just to see if students did something).

A large majority of teachers asked students to clarify their thinking or justify responses out loud, and asked questions at a variety of levels such as recall, comprehension, and inference (for both indicators, 86.7% at Time One; 77.3% at Time Two).

About one half of teachers used dipsticking methods or every-pupil-responds methods (57.8% of classes at Time One; 47.7% of classes at Time Two). Similar proportions called students up front to solve problems (53.3% at Time One; 45.5% at Time Two).

Exit cards or summarizers were used in about one third of classes at Time One (31.1%) and one fifth of classes at Time Two (20.5%). (Among classes where observers saw the entire class, these figures were 42.8% at Time One and 28.5% at Time Two.) Teachers listened to student discussions in pairs or groups in one fifth of classes (22.2% at Time One; 20.5% at Time Two).

Teacher use of interactive technology. Interactive classroom technology used in the observed classes was limited. The only somewhat common application of classroom technology was teachers using the Promethean board interactively, by calling up students to solve problems or uncover correct answers at the board (62.2% of Time One classes, 50% of Time Two classes).

Table 1 below summarizes the extent of evidence of each component in the observed classes.

Table 1. Delivery of Key Components of Math 7 Instruction

Practices	Extent of evidence[a] (N = 89 classes)
Classroom structures that support learning	
Students know what to do when they come into class; class group rules or expectations are posted.	high
Students can drop off completed work, get homework without teacher's help.	moderate
Students can get texts or calculators without teacher's help.	low
Critical thinking and questioning	
Teacher asks questions that focus on problem-solving strategies and reasoning, teacher models thinking process, teacher reinforces students' use of the language of mathematics, teacher helps students make connections to prior knowledge, teacher presents or demonstrates multiple strategies.	high
Teacher uses "real world" applications of mathematical concepts.	moderate
Discourse and group work	
Teacher has students work in small groups or pairs.	moderate
Teacher facilitates student discussions about mathematical concepts and processes; teacher has students discuss in groups or pairs.	low
Differentiation, variety, and learning styles	
Teacher uses a variety of materials and modalities to teach the lesson.	high
Teacher encourages students to try a variety of materials and methods.	moderate
Teacher varies activities, formats, or outcomes for individual students; teacher has students use strategies or seek resources other than the teacher; teacher gives students opportunities to make choices; teacher provides differentiated activities, formats, or outcomes for different groups of students.	low
Formative assessment	

Table 1. Delivery of Key Components of Math 7 Instruction

Practices	Extent of evidence[a] (*N* = 89 classes)
Asking direct questions to check for understanding, walking around and visually checking students' work, asking student to clarify thinking or justify response aloud, asking questions at a variety of levels	high
Every pupil responds/dipsticking; calling students to front of class to solve a problem	moderate
Exit card/summarizer; listens to students discussing in groups or pairs	low
Interactive technology	
Promethean board used interactively so that students participate	moderate

[a] High = six out of ten classes or more. Moderate = about four or five out of ten classes. Low = fewer than four out of ten classes.

Summative Study

The analysis discussed in this section included 87 teachers. This total included 44 observed teachers (see previous section) and 43 teachers not eligible for observation (see Method).

The students used for the analyses described in this section met additional criteria for completion of the course and availability of complete data records. There were 3,562 students who completed Math 7 for Grade 7. This included 966 students in an observed class section; 1,036 students with an observed teacher who were not in an observed class section; and 1,560 students of non-observed teachers.

Are recommended instructional practices used by Math 7 teachers significantly related to student outcomes, as measured by MSA mathematics? Multiple regression procedures were applied to examine the relationship between teachers' use of instructional strategies in Math 7 lessons and students' mathematics performance. The procedures were done separately for each factor as well as individual dichotomous indicators of the practice as identified in the methodology. The outcome measure (or the dependent variable) for this study question was the Grade 7 MSA mathematics score. The expectation was that better mathematic performance would be associated with the higher use of the recommended instructional strategies or indicators of the Math 7 instructional practices.

Classroom structures. Four dichotomous indicators were used to measure classroom structures that support learning in Math 7 (Table 2.1). The analyses revealed that the presence of two indicators in Math 7 classrooms were significant predictors *($p < .05$)* of students' MSA test scores. These indicators were: "Students appear to know what to do when they come into the room (e.g., find their seat, pick up work at front table) or when they form groups (e.g., find partners, move into groups)," and "Students can drop off completed work and get copies of homework or make-up work without teacher's help."

Similar analyses for the following two indicators of classroom structures showed no significant effect (positive or negative) of these indicators on students' MSA mathematics scores: "Class ground rules or expectations are posted," and "Students can get textbooks, calculators, active egg, etc. without teacher's help."

Table 2.1. Multiple Regression Results for Classroom Structures That Support Learning

	B	Std. error	*t*	Degrees of Freedom*	*p*
Indicators					
1. Class ground rules or expectations are posted.	1.335	0.958	139 3	1,718	0.164
2. Students appear to know what to do when they come into the room (e.g., find their seat, pick up work at front table) or when they form groups (e.g., find partners, move into groups).	8.895	0.979	9.08 9	1,718	0.000**
3. Students can get textbooks, calculators, active egg, etc. without teacher's help.	0.012	0.791	0.01 5	1,718	0.988
4. Students can drop off completed work and get copies of homework or make-up work without teacher's help.	2.26	0.80	2.81 5	1,718	0.005**

*The degrees of freedom for these analyses are N- # covariates- 1
**Significant in positive direction

Critical thinking. The principal component analyses revealed that two factors underlie the six observation indicators measuring critical thinking and questioning in Math 7 lessons. The first factor explained the highest proportion of the variance associated with the six indicators, and hence, is the most informative one.

The results from multiple regression showed that the first factor was significantly *($p < .05$)* associated with students' performance in MSA mathematics (Table 2.2). The following three indicators had high loadings on the significant factor: "Teacher asks students questions that focus on problem solving strategies and reasoning," "Teacher models thinking process for developing strategies and discovering relationships," and "Teacher reinforces students' use of the language of mathematics (vocabulary, speaking and writing)." This finding suggests that teachers' higher use of the above stated indicators in Math 7 lessons was significantly associated with students' higher scores on Grade 7 MSA mathematics.

The second factor of practice was not significant *(p < .05)*, suggesting the three indicators of practice loaded on this factor (Table 2.2) do not influence the students' MSA mathematic scores. The nonsignificant factor had high loadings for the following: "Teacher uses 'real world' applications of mathematical concepts," "Teacher presents or demonstrates multiple strategies to students," and "Teacher helps students make connections to prior knowledge."

Table 2.2. Multiple Regression Results for Critical Thinking and Questioning

Indicators	B	Standard error	*t*	Degrees of freedom *	*p*
Factor 1 1. The teacher asks students questions that focus on problem solving strategies and reasoning. 2. Teacher models thinking process for developing strategies and discovering relationships. 3. Teacher reinforces students' use of language of mathematics (vocabulary, speaking and writing).	1.08	0.396	2.73	1,718	0.006 **
Factor 2 1. Teacher uses "real world" applications of mathematical concepts. 2. Teacher presents or demonstrates multiple strategies to students. 3. Teacher helps students make connections to prior knowledge.	-0.543	0.397	-1.37	1,718	0.17

* The degrees of freedom for these analyses are: N- # covariates- 1
** Statistically significant in the positive direction

Discourse and group work. Three observation indicators measuring discourse and group work in Math 7 classes were dichotomous and therefore included separately in the multiple regression models. The findings (Table 2.3) reveal significant and positive effects *($p < .05$)* of the presence of one of the three indicators in Math 7 classes on the students' MSA test scores: "Teacher has students discuss in groups or pairs (turn to a partner or think pair share)."

Similar analyses found a significant negative association between the presence of the following indicator and students' MSA scores: "Teacher has students work in small groups or pairs to solve problems." Finally, there was no significant association between the indicator "Teacher facilitates student discussions about mathematical concepts and processes" and students' MSA mathematics scores.

Table 2.3. Multiple Regression Results for Discourse and Group Work

Indicators	B	Std. error	*t*	Degrees of freedom*	*p*
1. Teacher facilitates student discussions about mathematical concepts and processes.	-.047	0.998	-0.047	1,718	0.96
2. Teacher has students discuss in groups or pairs ("turn to a partner" or "think pair share").	2.388	0.935	2.555	1,718	0.011**
3. Teacher has students work in small groups or pairs to solve problems.	-5.52	0.87	-6.34	1,718	.000***

*The degrees of freedom for these analyses are N- # covariates- 1
** Significant in positive direction
*** Significant in negative direction

Differentiation, variety, learning styles. Five observation indicators were used to measure differentiation in Math 7 lessons. The principal component procedures found a three-factor solution from the five indicators. Further analyses (multiple regression) revealed that nether factor 1 (t=.05; df=1,718); p=0.96), factor 2 (t=0.98; df=1718) nor factor 3 (t=-0.56; df=1,718; p=.58) had a significant effect on students' mathematics performance as measured by Grade 7 MSA mathematics.

Formative assessment. The principle component procedures revealed a three-factor solution (Table 2.4) from the six indicators of formative assessment. The first factor had high loadings for four of the indicators and was negatively associated with the students' performance as measured by Grade 7 MSA mathematics. This finding implies that the higher use of the following indicators of the practice in Math 7 lessons was associated with lower students' MSA mathematics scores: "Asking direct questions to check for understanding and listening to students' responses," "Visual walk-around and check of homework or work at students' desks," "Every pupil responds/dipsticking/thumbs up," and "Calls students to front of class to solve a problem."

The second factor was also negatively associated with the students' mathematics performance, suggesting the teachers' higher use of listening to students discussing in pairs or groups is associated with lower students' scores on MSA mathematics.

The third factor, however, had a significant positive association with students' MSA mathematics. This suggests the use of the following strategy was significantly related to Grade 7 mathematics performance: "Asking student to clarify thinking or justify response aloud (critical thinking)."

Table 2.4. Multiple Regression Results for Formative Assessment

Indicators	B	Std. error	*t*	Degrees of freedom*	*p*
Factor 1	-1.17	0.394	-2.97	1,718	0.003**
1. Asking direct questions to check for understanding and listening to students' responses					
2. Visual walk-around and check of homework or work at students' desks (for content, not just that students did something)					
3. Every pupil responds/dispsticking/thumbs up					
4. Calls students to front of class to solve a problem					
Factor 2	-0.847	0.394	-2.148	1,718	0.032**
1. Listens to students discussing in pairs or groups					
Factor 3	0.859	0.401	2.14	1,718	0.032***
1. Asking student to clarify thinking or justify response aloud (critical thinking)					

*The degrees of freedom for these analyses are N- # covariates- 1
** Significant in negative direction
*** Significant in positive direction

Technology. Four indicators constituted the measurement system for the use of technology in Math 7 classrooms. These indicators measured teachers' use of Smart View (display) calculator with Promethean board or overhead projector, students' use of calculators as tools for understanding concepts, teachers' use of Internet tools to enhance instruction, and teachers' use of Promethean board interactively so that students participate. Two factors were extracted from the four indicators and were further included in the multiple regression analyses. The findings revealed that both factor 1 ((t=-1.9; df=1,718) and factor 2 (t=-0.14; df=1,718) were unrelated *(p < .05)* to Math 7 students' achievement as measured by MSA mathematics.

Other indicators. The analysis found that the use of an exit card or summarizer in Math 7 lessons was significantly associated *(p < .05)* with Math 7 students' MSA test scores (t=2.96; df=1,7188). The same analysis did not show a significant relationship (t=0.13; df=1718; p=0.90) between Math 7 students' test scores and the presence of teachers' use of the indicator: "Asking questions at a variety of levels (recall, comprehension, inference)."

Are there differences in mathematics performance between students of observed teachers and students of non-observed teachers? Eligible teachers for observation ("experienced teachers") were teaching Math 7 for Grade 7 for at least the second time in the past three school years (2010–2011, plus 2009–2010 and/or 2008–2009). There were 45 eligible teachers in 32 MCPS middle schools. These teachers are referred to as "observed teachers" below. Analysis of Covariance (ANCOVA) was used to statistically control for the effects of possible pre-existing differences between the two groups of students (students of observed teachers vs. students of non-observed teachers).

The results showed that the main effect of the teachers' experience was not significant (F = 0.615; df = 1; *p* < .05), after controlling for demographics, service receipt measures, and academic performance. The interaction effect[1] was not statistically significant at *p* = .05 level (F = .70; *p* = .40).[2] On average, students in the observed teachers' classroom scored (adjusted mean=402.6; n=1,723) the same as their peers (p>0.05; Standard Error=0.62) in the non-observed teachers' classrooms (adjusted mean=402.11; n=1,302). This finding is further confirmed by the calculated effect size (0.02).

[1] The product term between the independent variable (students of observed teachers vs. students of non-observed teachers) and the covariate (Grade 6 MSA) was included in the ANCOVA model. The coefficient of this product was used to test for non-parallelism or interaction.

[2] Levene's test for the equality of error variances between the two groups of students found that the variances were not significantly different (*F* = 1.64, *p* < .05).

Summary/Conclusions

Are Math 7 teachers of Grade 7 students using recommended instructional practices?

Classroom structures that support learning. Extent of implementation was found to be high for two indicators of structures: 1) Students appear to know what to do when they come into the room or move into groups. 2) Class ground rules or expectations are posted. Evidence for the indicator, "Students can drop off completed work and get homework without the teacher's help," was moderate. The extent of evidence for the indicator, "Students can get texts, calculators . . . without teacher's help," was low.

Critical thinking and questioning. Extent of implementation was found to be high for the following indicators: 1) Teacher asks students questions that focus on problem-solving strategies and reasoning. 2) Teacher models thinking process. 3) Teacher reinforces students' use of the language of mathematics. 4) Teacher helps students make connections to prior knowledge. 5) Teacher presents or demonstrate multiple strategies. The indicator "Teacher uses 'real world' applications of mathematical concepts" was at a moderate level of implementation.

Discourse and group work. Extent of implementation was at a moderate level for the indicator of "teacher has students work in small groups or pairs to solve problems." However, the extent of evidence was low for two other indicators: 1) Teacher facilitates student discussions. 2) Teacher has students discuss in groups or pairs.

Differentiation, variety, and learning styles. Extent of implementation was high for the indicator, "Teacher uses of a variety of materials and modalities to teach the lesson to the whole class." A moderate level of implementation was evident for the indicator "Teacher encourages students to try a variety of materials and methods."

Implementation was at a low level for the following indicators: 1) Teacher differentiates activities, formats, or outcomes for different groups of students. 2) Teacher has students use strategies or seek resources other than getting information from the teacher. 3) Teacher gives students opportunities to make choices.

Formative assessment. Implementation was found to be at a high level for the following indicators of formative assessment: 1) Asking direct questions to check for understanding. 2) Visual walk-around and check of work at students' desks. 3) Asking student to clarify thinking or justify responses aloud. 4) Asking questions at a variety of levels (recall, comprehension, inference). Implementation was at a moderate level for the indicators for dipsticking (every pupil responds) and for "Calls students to the front of class to solve a problem." Implementation was at a low level for using exit cards, and for "Listens to students discussing in pairs or groups."

Interactive technology. This analysis found a moderate level for teachers having students interact with the Promethean board. Other indicators of the use of interactive technology were implemented at a low level, including "Teacher has students use calculators to understand concepts," and "Teacher uses Internet tools to enhance instruction."

Are recommended instructional practices used by Math 7 teachers significantly related to student outcomes, as measured by MSA mathematics? After controlling for the students' prior ability (Grade 6 MSA) and characteristics, the multiple regression analyses found that several Math 7 instructional practices were significantly related to Grade 7 MSA mathematics scores. These significant (positive or negative) relationships are summarized below.

Classroom structures. The analyses revealed that the presence of two of the four indicators of classroom structure in Math 7 classrooms were positive significant predictors ($p < .05$) of students' MSA test scores. These indicators were: "Students appear to know what to do when they come into the room (e.g., find their seat, pick up work at front table) or when they form groups (e.g., find partners, move into groups)," and "Students can drop off completed work and get copies of homework or make-up work without teacher's help."

Critical thinking. Teachers' higher use of three instructional practices in Math 7 classrooms was positively and significantly ($p < .05$) associated with students' higher scores on MSA mathematics. These practices included: "Teacher asks students questions that focus on problem-solving strategies and reasoning," "Teacher models thinking process for developing strategies and discovering relationships," and "Teacher reinforces students' use of the language of mathematics (vocabulary, speaking and writing)."

Discourse and group work. The findings revealed significant and positive effects ($p < .05$) of the presence of the following indicator in Math 7 classrooms on the students' MSA test scores: "Teacher has students discuss in groups or pairs (turn to a partner or think pair share)." Similar analyses found that the presence of another practice placed in this category, "Teacher has students work in small groups or pairs to solve problems" in Math 7 classrooms was negatively associated with students' MSA mathematics score.

Formative assessment. The higher use of four instructional practices (loaded on the same factor) in Math 7 classrooms was significantly associated ($p < .05$) with lower students' mathematics scores as measured by MSA mathematics. These practices included: "Teacher asks direct questions to check for understanding and listening to students' responses," "Visual walk-around and check of homework or work at students' desks," "Every pupil responds/ dipsticking/thumbs up," and "Call students to front of class to solve problem." Another practice of formative assessment also was negatively associated ($p < .05$) with students' MSA mathematics scores, suggesting that the teachers' higher use of listening to students discussing in pairs or groups strategy is related with students' lower scores on MSA. Similar analyses found that the use of the following recommended practice was significantly and positively related ($p < .05$) to students' math performance as measured by

Grade 7 MSA mathematics: "Asking student to clarify thinking or justify response aloud (critical thinking)."

Other indicators. The analyses found that the use of an exit card or summarizer in the Math 7 classroom was positively significantly associated *($p < .05$)* with Math 7 students' MSA test scores.

Are there differences in mathematics performance between students of observed teachers and students of non-observed teachers? On average, students of observed teachers (those with more recent experience teaching Math 7 for Grade 7) performed as well as students of non-observed teachers, as measured by their Grade 7 MSA mathematics after controlling for students' initial abilities (Grade 6 MSA scores), demographics, and service receipt measures.

Recommendations

The following evidence-based recommendations were generated for school district guidance.

1. Enhance the use of those instructional practices in Math 7 classes that have been identified by this study to have positive and significant associations with MSA mathematics test scores.
2. Collaborate with the staff from the mathematics office to further improve the reliability and validity of the measures of Math 7 practice in the observation instrument.
3. Replicate the study over time, using different student populations and settings to see if the findings of this study are stable.

Strengths and Limitations Associated with the Study

In examining the effects of Math 7 classroom practices on students' MSA mathematics performance, it is important to consider that measuring instructional strategies in practice is not an easy task. In general, no measure is perfect, specifically measuring practices such as differentiation that needs concrete indicators and associated operational definitions. Therefore, achieving a psychometrically sound operationalization of Math 7 instructional practices is necessary when studying the impact of indicators of practice on the students' mathematics performance.

A number of steps were taken to safeguard a strong methodology and produce reliable results. 1) All Math 7 for Grade 7 teachers meeting the experience criteria were observed. 2) The inclusion of nearly all Math 7 for Grade 7 students, and the inclusion of all Math 7 for Grade 7 teachers, in the analysis allows findings to be generalized to all MCPS middle schools and ensures the external validity of results for the question of whether teacher experience affects performance. 3) The use of multiple data sources (observations, lesson logs, course materials and handouts, teacher information, and extensive student information) provides a more complete view of implementation and the current status of the Math 7 for Grade 7 course in MCPS. 4) The use of relevant literature and MCPS mathematics content experts to guide the development of indicators or measures of Math 7 practices ensures that the findings: a) provide a set of relevant measures and their psychometric properties that are useful for the future studies of Math 7 instructional practices, and b) identify those recommended practices in Math 7 lessons that are significantly associated with student performance on Grade 7 MSA mathematics. 5) Observations were conducted during brief and specific time periods (once at the beginning of Unit 2 and once at the end of Unit 2), strengthening the ability to assess the Math 7 environment at a specific point in time and at a specific point in the course curriculum. 6) Each eligible teacher was observed twice, teaching the same group of students. Multiple observations using the same protocol help to ensure that instructional practices not seen in one class may be seen in another.

The following limitations pertain to this study. 1) Isolating the effects of the recommended practices on students' test scores is not an easy task. There are many factors that can also affect students' test scores but could not be controlled in this study due to their unavailability. Only a classical experiment with the random assignment of students safeguards against each of the sources of internal invalidity in a study (e.g., selection bias, maturation, history, attrition). (Babbie, 1992; Judd, Smith, & Kidder, 1991). 2) MCPS has not done another study relating instructional practices to student performance, so a comparison to method or findings for any prior MCPS study is not available.

Finally, in drawing conclusions from this study, three caveats must be noted. 1) Generalization of the findings for relationship of teacher practice to student performance is limited to: a) the students of observed teachers and b) those students who took the Grade 7 MSA in mathematics. Students who took the Mod-MSA or Alt-MSA were excluded from analyses since the scale scores resulting from modified or alternative versions of the MSA cannot be combined with or compared with those from the standard MSA administration. 2) Only one criterion of students' mathematics achievement which was available was used to examine the relation between recommended Math 7 practices and students' mathematics performances. The relation may change if, for example, unit assessments were used as the outcome measure in the analyses. In addition, other outcomes of recommended Math 7 practices (e.g., having better aptitude in mathematics learning) were not addressed in this study. 3) Although the findings obtained from this study were based on sound design as well as appropriate analyses, it should be noted that causality may not be inferred from this study due to the lack of an experimental design. The outcomes of the instructional practices, whether measured in terms of MSA or other tests, depend on a complex set of interactive factors that can be better addressed by a randomized study.

A technical report featuring project background, methodology, detailed findings, statistical procedures, and recommendations is available at: http://sharedaccountability.mcpsmd.org/reports/list.php?selection=905.

References

Earl Babbie, *The practice of social research*, 6th Edition (Belmont, CA: Wadsworth, 1992).
Rachel A. Hickson, *Evaluation study: Preparing students for Algebra 2* (Rockville, MD: Montgomery County Public Schools, 2010).
C. M. Judd, E. R. Smith, and L. H. Kidder, *Research methods in social relations*. (San Francisco: Holt, Rinehart and Winston, Inc., 1991).
J. K. Luellen, W. R. Shadish, and M. H. Clark, "Propensity scores: An introduction and experimental test," *Evaluation Review*, Vol. 29, No. 6 (2005): 530–558.
Montgomery County Public Schools, *Mathematics instructional guide, Math B,* Draft (Rockville, MD: Author, 2003).
Montgomery County Public Schools, *Curriculum quick reference, mathematics, Math 7/ Math B,* Draft (Rockville, MD: Author, 2007).
P. R. Rosenbaum and D. B. Rubin, "The central role of the propensity score in observational studies for causal effects," *Biometrika* 70 (1983): 41–45.
P. R. Rosenbaum and D. B. Rubin, "Reducing bias in observational studies using subclassification on the propensity score," *Journal of the American Statistical Association* 79 (1984), 561–524.
P. R. Rosenbaum and D. B. Rubin, (1985), "Constructing a control group using multivariate matched sampling that incorporates the propensity score," *The American Statistician* 39 (1985): 33–38.
W. R. Shadish, T. D. Cook, and D. T. Campbell, *Experimental and quasi-experimental designs for generalized causal inference* (Boston, MA: Houghton Mifflin Company, 2002).
U.S. Department of Education (USDE), Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, What Works Clearinghouse. *Procedures and standards handbook:* Author, 2008.