MFAS 2012-13 Year-Long Study Results

Laura Lang, Principal Investigator, (llang@fsu.edu)
Mark LaVenia, Methodologist, (mlavenia@lsi.fsu.edu)

For more information about MFAS, please go to: http://mfas.cpalms.org

Abstract

The purpose of this report is to relay findings from the Mathematics Formative Assessment System (MFAS) 2012-13 study conducted with Kindergarten and Grade 1 teachers and students. The study was a year-long cluster randomized field trial, conducted August 2012 through May 2013. Schools were randomly assigned to either the MFAS treatment condition or a business-as-usual control condition. Positive statistically significant effects were found for MFAS on teacher mathematics knowledge for teaching, classroom instructional practice, and student mathematics performance. Namely, treatment teachers scored on average more than a quarter of a standard deviation higher than control teachers on the math knowledge posttest. Classroom observation data revealed that the odds of observing small group instruction were more than 500 times greater for treatment classrooms compared to control classrooms and the odds of observing whole group instruction as the predominant grouping strategy were .87 times lower for treatment classrooms compared to control classrooms. Lastly, mean treatment group student math performance at posttest was nearly a third of a standard deviation higher than mean control group student math performance. The effect sizes on Kindergarten and Grade 1 student mathematics performance were $g = .32$ and $g = .29$, respectively. Findings from principal interviews corroborated these results, where principals reported observing changes in teachers' instructional practices as a result of participation in MFAS. These changes were universally seen as improvements to teacher knowledge and instructional practice. Examples of changes noted by principals included more individualized, differentiated instruction; a paradigm shift for teachers, especially veteran teachers; and higher expectations for student thinking and performance. Effects found for MFAS on teacher mathematics knowledge and students' mathematics performance were consistent with results we have found in prior pilot studies conducted on MFAS. The benefits for teachers to engage in communities of instructional practices to discuss student work and analyze formative assessment data for the purposes of differentiated instruction are supported in these findings. Further, these findings also support that such practices by teachers have demonstrable benefits for their students.

Questions regarding the MFAS research and analyses should be directed to Laura Lang, principal investigator, (llang@fsu.edu) or Mark LaVenia, methodologist, (mlavenia@lsi.fsu.edu). This study was funded by the Florida Department of Education, award numbers 371-RA211-4C001 and 371-RA411-4C001. The opinions expressed are ours and do not represent views of the funding agency.

Background

Three randomized field trials have been conducted on the Mathematics Formative Assessment System (MFAS): a nine-week pilot study in the spring of SY 2009-10 with Grades K-3 teacher teams using tasks aligned with Florida's Next Generation Sunshine State Standards (NGSSS); a semester-long pilot study in the winter and spring of SY 2011-12 with Grades 2-3 teacher teams using tasks aligned with Florida's NGSSS; and a year-long study in SY 2012-13 with Grades K-1 teacher teams using tasks aligned with the Common Core State Standards (CCSS). The 2010 nine-week study was conducted with support from Florida Department of Education funds. The 2012 semester-long and 2012-13 year-long studies were conducted through a competitively awarded grant from Florida's Race to the Top funding. The 2012-13 year-long study made use of intervention protocols and measures developed, field tested, and validated in the prior two pilot studies (e.g., Schoen, LaVenia, Lang, & Oberlin, 2013). In addition, results from each successive study were applied iteratively to make programmatic improvements as warranted by study findings.

Findings from the MFAS 2010 and 2012 pilot studies have culminated in the theory of action illustrated in Figure 1. The MFAS theory of action posits that the use of MFAS will lead to changes in teacher knowledge, skills, and beliefs, which will iteratively continue to develop as a result of the use of differentiated instruction supported by math formative assessment. The continued development of teacher knowledge, skills, and beliefs resulting from improved implementation of differentiated instruction would then lead to increased student mathematics achievement. The MFAS theory of action accounts for the reality that district and school contextual variables, teacher characteristics, and student characteristics represent important input to this process.

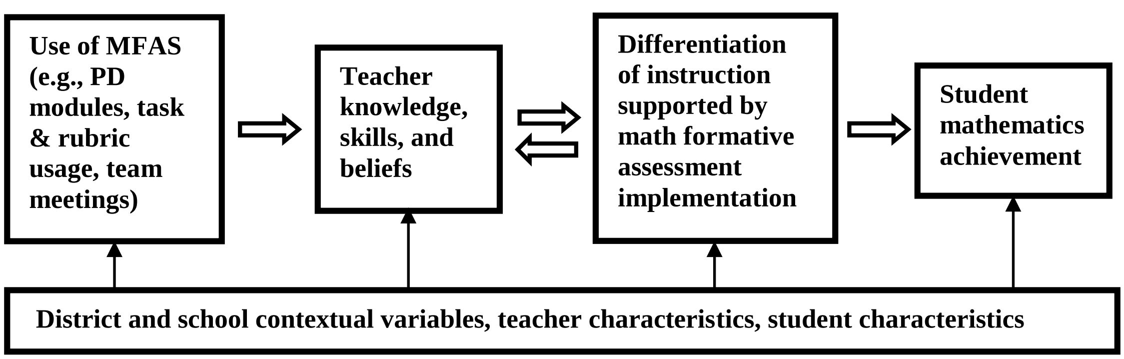

Figure 1. **MFAS Theory of Action**

Key Findings from the MFAS 2010 Nine-Week Pilot Study. The 2010 nine-week pilot study was a cluster randomized field trial with 169 teachers across grade levels K-3 from 14 schools. The focus of the study was on a small sample of MFAS tasks and rubrics, combined with a half-day of initial training and loosely structured grade-level team meetings (Lang, Hawthorne, Sakon, Reta, & Schoen, 2011). Results indicated that when fully implemented, MFAS had a statistically significant positive effect on the end-of-semester NGSSS math benchmark test (d = 0.62). Many of the treatment group teachers initially rejected the tasks because they believed the tasks to be too difficult for students. After implementing the tasks and seeing that their students learned to successfully solve the problems when instruction targeted the gaps in their knowledge, these teachers reported major changes in beliefs and increased expectations for students in mathematics (Lang, Schoen, Davis & Howell, 2010). Although this initial feasibility trial indicated the potential of MFAS to improve elementary mathematics teaching and learning, a key finding was that many teachers required much more extensive professional development and ongoing support in order to incorporate formative assessment in their day-to-day practice.

Key Findings from the MFAS 2012 Semester-Long Pilot Study. The 2012 semester-long study was a cluster randomized field trial with 160 teachers across grade levels 2-3 from 21 schools in two districts. The study investigated the use of teacher-selected MFAS tasks and rubrics, combined with the three online professional development (PD) modules and structured face-to-face weekly team meetings facilitated by project staff. Analyses suggest positive effects for those teachers with relatively higher levels of implementation. Specifically, students of teachers who implemented five or six tasks related to the tested benchmarks (the upper quartile for task implementation) scored significantly higher than control students on the NGSSS math benchmark test. On average, we estimated an approximate effect size of 0.6 between high-level use and nonuse (d =.7 in Grade 2 and d =.55 in Grade 3). This finding is consistent with what we found in our prior nine-week pilot. Findings suggesting that level of implementation moderates the impact of MFAS highlight the importance of related teacher professional learning and ongoing support during implementation.

Methods for the MFAS 2012-13 Year-Long Study

Setting and sample. The MFAS 2012-13 year-long study was conducted in three school districts: a medium-sized rural district located in the Florida Panhandle (District A); a medium-sized suburban district located in the central region of Florida (District B); and a large urban district also located in the central region of Florida (District C). Thirty-two schools across the three districts (six in District A, 14 in District B, and 12 in District C) were recruited for participation. One school in District B withdrew from the study during the first quarter of the school year, resulting in a final sample of 31 schools. The final sample comprised 301 consenting Kindergarten and Grade 1 teachers (Kindergarten n = 146; Grade 1 n = 155). Parent consent to participate in the study was obtained for 2,317 Kindergarten and 2,515 Grade 1 students. Teacher teams were comprised of three or more teachers with a common instructional assignment (i.e., grade level). Due to missing data the three-district analytic sample of students is reduced to Kindergarten n = 2048 and Grade 1 n = 2222. Further, some data on student background characteristics (i.e., Free/Reduced-price Lunch status and English Language Learner status) were missing entirely from District A; thus for analyses that include those variables, the analytic sample of students was further reduced to Kindergarten n = 1613 and Grade 1 n = 1779. Tables 1 and 2 relay sample size disaggregated by student background characteristic for Kindergarten and Grade 1, respectively.

Table 1.
Sample Size by Student Background Characteristics Disaggregated by Grade Level

Characteristics	Treatment	Control	Total
		Kindergarten	
Free/Reduced-price Lunch (out of *n* = 1613)	433	532	965
Race/Ethnicity (out of *n* = 2048)			
Asian, Non-Hispanic	44	38	82
Black, Non-Hispanic	166	151	317
Hispanic or Latino	184	204	388
American Indian, Non-Hispanic	3	2	5
Multiracial, Non-Hispanic	56	50	106
White, Non-Hispanic	513	637	1150
English Language Learners (out of *n* = 1613)	51	72	123
Students With Disabilities (out of *n* = 2048)	98	112	210
		Grade 1	
Free/Reduced-price Lunch (out of *n* = 1779)	523	545	1068
Race/Ethnicity (out of *n* = 2222)			
Asian, Non-Hispanic	62	55	117
Black, Non-Hispanic	190	154	344
Hispanic or Latino	224	202	426
American Indian, Non-Hispanic	0	4	4
Multiracial, Non-Hispanic	57	41	98
White, Non-Hispanic	584	649	1233
English Language Learners (out of *n* = 1779)	89	98	187
Students With Disabilities (out of *n* = 2222)	121	107	228

Note. The reduced sample size for FRL and ELL is due to missing data on those variables.

Random assignment of schools. Randomization was blocked on district and percent Free/Reduced-price Lunch (FRL), resulting in equal proportions of treatment and control schools within each district and near equal mean percentage of FRL across the two conditions.

Intervention. The purpose of the MFAS intervention was to increase (a) teacher knowledge of the learning goals defined by the mathematics content in the Common Core State Standards, (b) teacher use of formative assessment to guide instruction, (c) teacher collaboration based around evidence of students' mathematical thinking, and (d) differentiation of instruction in mathematics based upon the instructional needs of individual students.

Teachers in the treatment group (a) completed three introductory professional development modules available online through CPALMS, (b) implemented MFAS tasks related to the standards they were teaching, (c) used MFAS rubrics to evaluate student performance and differentiate instruction, and (d) participated in weekly community of instructional practice (CIP) meetings with members of their grade-level team. Facilitated by district lead teachers, the CIP meetings focused on implementation of MFAS-CCSS and differentiation of mathematics instruction based on the results. Also, principals of treatment schools completed an online module for school leaders and committed to support their teachers in implementing the MFAS tasks in their classrooms and collaborating within their teams in using the MFAS system.

Similar to all teachers in Florida, teachers and principals in the control group had access to the MFAS professional development modules as well as tasks and rubrics through the CPALMS website. However, they were not expected to complete the modules, use the MFAS tasks and rubrics to differentiate instruction, or participate in weekly CIP meetings with grade-level teams.

Incentivization. Schools were recruited in the late spring and early summer of 2012. In exchange for participation, teachers received all materials and training at no cost. The treatment group received training on the MFAS-CCSS tasks and rubrics through embedded, on-the-job training, including weekly meetings with their communities of instructional practice. As compensation for additional teacher time and effort, each participating teacher, principal, and assistant principal at schools assigned to the MFAS condition received a $500 stipend or gift card and each participating teacher and principal at schools assigned to the non-MFAS condition received a $300 stipend or gift card. The form of the incentive, stipend or gift card, was based on each participant's preference.

Measures. The measures used in this study include the following: the Learning Mathematics for Teaching (LMT; Hill, Schilling, & Ball, 2004) assessment of teacher knowledge for teaching elementary grades mathematics; the FCR-STEM Observation Protocol for Formative Assessment in Mathematics (OPFAM); and the FCR-STEM CCSS Student Mathematics Assessments for Kindergarten and Grade 1.

Data collection schedule. All participating teachers completed pre and postforms of the LMT, in August and May, respectively, of SY 2012-13. A subsample of two teachers per grade level per school was randomly selected for classroom observation. Trained observers conducted the classroom observations the first two weeks of April 2013. Trained proctors administered student assessments to all participating students in MFAS and non-MFAS classrooms in May 2013. Also, a sample of eight principals (from across the three districts) leading MFAS treatment group schools were interviewed in May 2013 by Laura Lang, Principal Investigator.

Analytic strategy. Data were modeled using a multilevel mixed model procedure with random effects for teacher and school. Thus, LMT and OPFAM scores were analyzed with a two-level model (teachers nested within schools) and student mathematics assessment data were analyzed with a three-level model (students nested within teachers nested within schools). Analyses of the impact of MFAS on LMT scores controlled for LMT scores at pretest. Analyses of the impact of MFAS on student math performance scores controlled for assignment block, student race/ethnicity, and disability status. For calculation of effect sizes associated with categorical independent variables (e.g., treatment condition), we used a Hedges' *g* weighted pooled standard deviation in the denominator to calculate unbiased mean group differences. For effect sizes associated with continuous independent variables (e.g., pretest score), we used a partial eta squared correlation ratio to estimate the proportion of variance in Y explained by X, after excluding variance explained by other predictors.

Findings

Presented below are findings for the main effects of MFAS on teacher mathematics knowledge for teaching, teacher classroom instructional practice, and student mathematics performance (Lang, Schoen, LaVenia, & Oberlin, 2014). In addition, we present results for an exploratory post hoc analysis of the MFAS effect on student mathematics, disaggregated by student subgroup.

Teacher mathematics knowledge for teaching. Analyses of the LMT data revealed a positive statistically significant effect for MFAS on teacher mathematics knowledge for teaching. Table 2 relays effect sizes across three models: Model 1 relays the effect size and statistical significance when Treatment is the only predictor included in the model, and Models 2 and 3 relay the treatment effect after controlling for pretest and grade level of teacher. Pretest was revealed to be a strong predictor, accounting for approximately two thirds of the variability on the posttest. However, no group difference between grade levels was observed, indicating that knowledge levels for the Kindergarten and Grade 1 teachers were indistinguishable. The conditional effect size for Treatment was estimated at $g = .27$. Thus, holding all other variables constant, MFAS teachers scored on average more than a quarter of a standard deviation higher than control teachers on the LMT posttest.

Table 2.
Effect of MFAS on Teacher Mathematics Knowledge for Teaching

	Model 1 Effect size	**Model 2 Effect size**	**Model 3 Effect size**
School-level			
Treatment	0.35*	0.27*	0.27*
Teacher-level			
LMT Pretest		0.67***	0.67**
Grade 1			–0.07

Note. Teacher $n = 301$ ($n = 156$ Control; $n = 145$ Treatment). Unconditional between-school ICC is $\rho = .07$.
*$p < .05$. **$p < .01$. ***$p < .001$.

Classroom observation. Statistically significant treatment effects were found for use of student grouping strategies. Classroom observers rated teachers on (a) whether Whole group, Small group, or One-on-one occurred at all and (b) which grouping strategy was the predominant strategy used by the teacher. Table 3 relays results which suggest that the odds of observing small group instruction were far greater in the treatment classrooms, with the odds ratio of $OR = 506$ indicating that the odds of observing small group instruction were more than 500 times greater for treatment classrooms compared to control classrooms.[1] In addition, observation

[1]Although the magnitude of this estimate may be distorted somewhat by the low prevalence of small group instruction in control classrooms, the result remains that there was a very large statistically significant difference between treatment and control classrooms with regard to the occurrence of small group instruction.

results suggest that the use of whole group as the predominant grouping strategy was considerably lower for the treatment group, with the odds ratio of *OR* = 0.13 indicating that the odds of observing whole group instruction as the predominant grouping strategy were 0.87 times lower for treatment classrooms compared to control classrooms.

Table 3.
Effect of MFAS on Teacher Use of Student Grouping Strategies

	Log odds (SE)	**Odds ratio**
Rated strategy as having occurred		
Whole group	0.56 (5.55)	1.75
Small group	6.23 (1.44)	506.03***
One-on-one	1.26 (0.69)	3.52˜
Rated as predominant strategy used		
Whole group	–2.04 (0.92)	0.13*
Small group	1.09 (1.62)	2.98
One-on-one	2.30 (2.67)	9.96

Note. Teachers *n* = 122 (*n* = 59 Treatment; *n* = 63 Control). Solitary observers conducted observations for 28 classrooms; rater-pairs conducted observations for 47 classrooms (Treatment *n* = 24 and Control *n* = 23) classrooms, resulting *n* = 169 Observations. The analytic model accounted for repeated measurement of classrooms by rater-pairs
˜$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

Student math performance. Analyses of the student mathematics assessment data reveal a positive statistically significant effect for MFAS on student mathematics. Tables 4 and 5 relay effect sizes for MFAS on student math performance at Kindergarten and Grade 1, respectively. Each table includes three models: Model 1 includes treatment status as the only predictor; Model 2 is Model 1 plus covariates for student background characteristics; and Model 3 is Model 2 plus controls for the assignment blocks. Student race/ethnicity and disability status are coded with White and Non-Disability as the reference categories, respectively.[2] Assignment blocks were based on district and school percent FRL; thus, inclusion of assignment blocks in Model 3, by proxy, accounts for between-school variation in the outcome explained by district and percent FRL. Coefficients for block are not presented in the table to save space, but are indicated in Model 3 by a vertical ellipse (⁝).

Effects of MFAS on Kindergarten student math performance. Average annual gains in effect size for math nationally-normed tests suggest an expected trajectory of 1.14 standard deviation growth from Kindergarten to Grade 1 (Bloom, Hill, Black, & Lipsey, 2008). The observed MFAS effect of 0.32 on Kindergarten mathematics represents an added gain of more than a fourth (0.32/1.14 = 0.28) of a year’s worth of growth—an additional 10 weeks of learning—for using MFAS. To help judge the practical importance of a 0.32 effect size, we translate the effect

[2] Interaction terms are not included in these models. Thus, statistically significant effects for student characteristics do not indicate anything about whether MFAS was more or less effective for a given group. Rather, effects for student characteristics for these models signify differences in the outcome between a given category (e.g., Disability) and the reference category (e.g., Non-Disability), on average, across the sample.

size into an improvement index that indicates the expected change in percentile rank for an average comparison group student, had the student received the intervention. An effect size of 0.32 translates to an improvement index of 13 percentile points. Accordingly, we can conclude that the intervention would have led to a 13% increase in percentile rank for an average student in the comparison group and that 63% (13% + 50%) of the students in the intervention group scored above the comparison group mean.

Table 4.
Effect of MFAS on Kindergarten Student Mathematics Performance

Fixed Effects	Model 1 Effect size	Model 2 Effect size	Model 3 Effect size
School-level			
Treatment	0.26~	0.29*	0.32**
Block			⁝
Student-level			
Asian		0.33**	0.33**
Black		–0.51***	–0.49***
Hispanic		–0.32**	–0.28**
American Indian		–0.14	–0.17
Multiracial		0.09	0.09
Disability		–0.36***	–0.35***

Note. Student n = 2048 (Treatment n = 966; Control n = 1082). Unconditional between-teacher ICC is ρ = .06 Unconditional between-school ICC is ρ =.10. Coefficients for assignment blocks are not presented to save space, but indicated in the table by a vertical ellipse (⁝).
~p < .10. *p < .05. **p < .01. ***p < .001.

Effects of MFAS on Grade 1 student math performance. Average annual gains in effect size for math nationally-normed tests suggest an expected trajectory of 1.03 standard deviation growth from Grade 1 to Grade 2 (Bloom et al., 2008). The observed MFAS effect of 0.29 on Grade 1 mathematics represents an added gain of more than a fourth (0.29/1.03 = 0.28) of a year's worth of growth—an additional 10 weeks of learning—for using MFAS. An effect size of 0.29 translates to an improvement index of 11 percentile points. Accordingly, we can conclude that the intervention would have led to an 11% increase in percentile rank for an average student in the comparison group and that 61% (11% + 50%) of the students in the intervention group scored above the comparison group mean.

Table 5.
Effect of MFAS on Grade 1 Student Mathematics Performance

Fixed Effects	Model 1 Effect size	Model 2 Effect size	Model 3 Effect size
School-level			
Treatment	0.29~	0.30*	0.29**
Block			⁝
Student-level			
Asian		0.22*	0.21*
Black		–0.47***	–0.45***
Hispanic		–0.40***	–0.33**
American Indian		0.70	0.73
Multiracial		–0.08	–0.07
Disability		–0.35***	–0.35***

Note. Student $n = 2222$ (Treatment $n = 1105$; Control $n = 1117$). Unconditional between-teacher ICC is $\rho = .07$ Unconditional between-school ICC is $\rho = .12$. Coefficients for assignment blocks are not presented to save space, but indicated in the table by a vertical ellipse (⁝).
~$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

Exploratory analyses of MFAS treatment effects on student mathematics by student subgroup. With regard to the post hoc subgroup analyses, there are two important caveats: (a) some of the subgroup sample sizes are small—resulting in underpowered analyses that suppress our ability to observe statistically significant effects and (b) testing across many subgroups involves multiple comparisons that increase our Type I error rate and chance for false discovery (i.e., an observed statistically significant effect that is found to be significant by chance and is not likely to replicate). Therefore, we employ the following strategy for interpreting results from the subgroup analyses. To address the underpowered issue, we believe that it is appropriate to consider p-values < .10 (rather than the conventional $p < .05$) as statistically significant. To address the issue of potential for false discover, we recommend reserving greatest confidence for only those effects that were observed in both grade levels (i.e., replicates across grade levels).

Table 6 relays the results from exploratory analyses of MFAS effects on student mathematics performance disaggregated by student subgroup. In Kindergarten, statistically significant positive effects of MFAS were found for Male and Female students; White students; Students with and without disabilities; FRL eligible and FRL noneligible students; and ELL and non-ELL students. Statistically significant effects in Kindergarten range from an effect size of $g = .20$ (Improvement Index of +8 percentile points) for Female students to an effect size of $g = .61$ (Improvement Index of +23 percentile points) for ELL students. In Grade 1, statistically significant positive effects of MFAS were found for Male and Female students; Hispanic, Multiracial, and White students; Students without disabilities; FRL eligible and FRL noneligible students; and non-ELL students. Statistically significant effects in Grade 1 range from an effect size of $g = .20$ (Improvement Index of +8 percentile points) for FRL-eligible students to an effect size of $g = .68$ (Improvement Index of +25 percentile points) for Multiracial students. Positive treatment effects that replicated across grade levels were found for the following student subgroups: Male and Female students; White students; Students without disabilities; FRL eligible and FRL noneligible students; and non-ELL students. There were no statistically significant negative effects; thus, for every subgroup, treatment group students performed as well or better than comparison group students.

Table 6.
MFAS Effects on Student Mathematics Performance Disaggregated by Student Subgroup

	Kindergarten			Grade 1	
	n	Effect size		*n*	Effect size
Overall effects					
Districts A, B, and C	2048	0.32**		2222	0.29**
Districts B and C[a]	1613	0.36**		1779	0.27*
Effects by subgroup					
Female	996	0.20*		1122	0.32**
Male	1052	0.30**		1100	0.22*
Asian	82	0.51		117	0.28
Black	317	–0.11		344	0.21
Hispanic	388	0.20		426	0.30**
American Indian	5	*insufficient n*		4	*insufficient n*
Multiracial	106	0.33		98	0.68*
White	1150	0.36*		1233	0.28**
Non-Disability	1838	0.25**		1994	0.28**
Disability	210	0.24˜		228	0.20
Non-FRL[a]	648	0.37**		711	0.37**
FRL[a]	965	0.21˜		1068	0.20˜
Non-ELL[a]	1490	0.27**		1592	0.26**
ELL[a]	123	0.61*		187	0.10

[a]Sample constrained to Districts B and C, due to missing data in District A.
˜$p < .10$. *$p < .05$. **$p < .01$.

MFAS Study Principal Interviews

Described below are the common themes in principal responses. Each of the principals responded to the same set of questions and provided the opportunity to give additional feedback relevant to the implementation of MFAS and/or the study.

Common themes expressed by principals.

- The MFAS Principal/Mathematics Coach Professional Development Module adequately prepared principals to support teacher implementation of MFAS. Principals described the PD as well organized and effective.
- The two suggestions for improving the MFAS Principal/Mathematics Coach PD Module were (1) adding additional video clips of successful implementation and (2) more in-depth coverage of topics related to the Common Core State Standards in Mathematics.

- The greatest challenge for principals in supporting teacher implementation of MFAS is the need for more mathematics instructional time in the daily schedule and adequate time for the weekly teacher meetings. Based on their reports to principals, this challenge is also a concern for teachers. Also, principals reported that teachers struggled to understand mathematics content demands of the CCSS.
- The principals overwhelmingly reported that the communities of practice (weekly teacher meetings by grade level supported by the MFAS district facilitator) provided teachers with the support for successful MFAS implementation. Nearly all the principals said they could not have been successful without the communities of practice.
- Principals reported observing changes in teachers' instructional practices as a result of participation in MFAS. These changes were universally seen as improvements to instruction and teacher knowledge. Some examples: more individualized, differentiated instruction; a paradigm shift for teachers, especially veteran teachers; and higher expectations for student thinking and performance.
- Principals reported seeing positive changes in student behavior that include improved student confidence, evidence of deeper thinking, awareness of learning goals and their progress vis-à-vis the criterion performance as well as more engagement and excitement about math.
- There was a very strong pattern in the comments that the importance of a trained MFAS facilitator meeting weekly with grade level teacher teams cannot be overstated. Most principals stated explicitly that without the facilitator, successful implementation of MFAS during year one would not have occurred to the extent that was observed.

Conclusions and Discussion

Effects found for MFAS on teacher math knowledge for teaching, classroom practice that is conducive to differentiated instruction, and students' mathematics achievement corroborated the results we have found in prior pilot studies. The benefits for teachers to engage in communities of instructional practices to discuss student work and analyze formative assessment data for the purposes of differentiated instruction are supported in these findings. Further, these findings also support that such teacher practices have demonstrable benefits for their students.

References

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness, 1*, 289-328.

Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal*, 105(1), 11.

Lang, L., Hawthorne, H., Sakon, F., Reta, A., & Schoen, R.C. (2011). Examining the effects of a K-3 mathematics formative assessment system. In J. Novotna & H. Moraova (Eds.), *International Symposium Elementary Maths Teaching: The Mathematical Knowledge Needed for Teaching in Elementary Schools* (pp. 214-221).

Lang, L., Schoen, R., Davis, D. & Howell, L. (2010, October). *Math K-3 Formative Assessment System for Florida's NGSSS*. Paper presented at the annual meeting of the Florida Council of Teachers of Mathematics, Orlando, FL.

Lang, L. B., Schoen, R. C., LaVenia, M., & Oberlin, M. F. (2014, March). *Mathematics Formative Assessment System - Common Core State Standards: A Randomized field trial in kindergarten and first grade.* Paper presented at the annual meeting of the Society for Research on Educational Effectiveness, Washington, DC.

Schoen, R. C., LaVenia, M., Lang, L. B., & Oberlin, M. F. (2013, April). *Developing a Valid, Reliable Observational Measure of Formative Assessment*. Paper presented at the annual meeting of the National Council of Teachers of Mathematics Research Presession, Denver, CO.