Examining Teacher Outcomes and Student Outcomes in the Math in the Middle (M^2) Institute Partnership

Stephen J. Meyer
John T. Sutton
RMC Research Corporation

2009 American Evaluation Association Conference
Orlando, FL
November 13, 2009

RMC RESEARCH

This Session

- Overview of M^2 and Evaluation Design
- Analysis of Teacher Outcomes
- Analysis of Student Outcomes
- Discussion of Limitations and Next Steps

RMC RESEARCH

M^2 Overview

- Partnership (U. of Nebraska, Lincoln; regional Educational Service Units; local school districts)
- Focus on
  - Middle-level teachers (G5-8)
  - Building capacity in rural settings

M^2 Program Components

- The M^2 Institute
  - 12 courses (2.5 years; intensive summer institutes and academic year courses)
  - Cohorts of approx. 30 teachers each year (most LPS in Y1)
  - Focus on developing deep mathematical content knowledge, pedagogy for middle level classrooms, action research, encouraging mathematical habits of mind, and leadership skills
- Mathematics Learning Teams
  - Lead teachers, supported by administrators and university faculty work with other teachers to improve instruction and assessment
- A Research Initiative
  - Additional examination of how M^2 components affect educational improvement and innovation

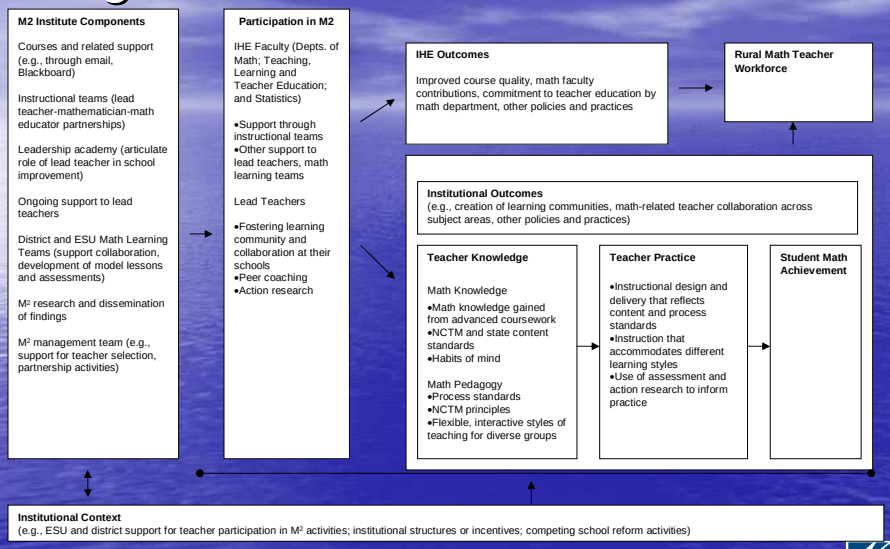
RMC Research Evaluation

- Quasi-experimental design; compares outcomes for M^2 participants and their students to those of teachers and students in a comparison group
- Data sources
  - Teacher surveys
  - Student achievement data
  - Content knowledge assessment
  - Interview and focus groups
  - Classroom observation
  - Document Analysis

Summary of Evaluation Questions

- Impact on student math achievement
- Quality of professional development activities
- Progress toward M^2 goals (related to teacher knowledge and practice, addressing different learning styles, action research, embedding math in other subject areas, developing learning communities & effective leaders)
- Impact on IHEs and IHE faculty practice
- Factors impeding or facilitating progress
- Progress toward "scaling up" and sustainability

Logic Model



Sample Findings from Several Publications

Sutton, J. T., Meyer, S., Brodersen, R. M., Jesse, D. and Northup, J. (2009, May). *2007-2008 Evaluation report: Math in the middle mathematics and science partnership program.* Denver, CO: RMC Research Corporation.

Meyer, S. J., & Sutton, J. T. (2008, March). *Examining teacher outcomes and student mathematics achievement outcomes in the Math in the Middle (M2) Institute Partnership*. Paper presented at the American Educational Research Association Annual Meeting, New York, NY.

Sutton, J.T., Meyer, S.J., Brodersen, R.M., and Turnbull, J.J. (2008). *University of Nebraska-Lincoln Math in the Middle Institute Partnership 2006-2007 evaluation report*. Denver, CO: RMC Research Corporation.

Sutton, J. T., Meyer, S., & Turnbull, J. (2007, March). *University of Nebraska–Lincoln. Math in the middle institute partnership: 2005–2006* (Evaluation report). Denver, CO: RMC Research Corporation.

Teacher Outcomes

Teacher Survey Measures

	Number of Items	Cronbach's Alpha
Overall Mathematics Professional Development Participation	**12**	**.83**
Overall Professional Development Emphasis on Mathematics Topics	**5**	**.86**
Overall Preparedness for Teaching Mathematics	**17**	**.91**
Preparedness to Teach Diverse Students (subscale)	4	.82
Overall Confidence in Mathematics and Teacher Support	**10**	**.90**
Confidence in Mathematical Knowledge (subscale)	3	.76
Confidence in Ability to Support Colleagues (subscale)	5	.88
Confidence in Leadership Ability (subscale)	2	.93
Emphasis on NCTM Process Standards	**5**	**.73**
Instructional Technology Use in Mathematics	**18**	**.83**
Overall Use of Assessment in Mathematics	**11**	**.78**
Use of Assessment - Analysis and Justification (subscale)	3	.82
Overall Factors that Limit Mathematics Teaching	**11**	**.70**
Factors that Limit Teaching - Student Characteristics (subscale)	5	.80
Factors that Limit Teaching - Instructional Resources (subscale)	4	.73
Overall Influence of External Factors on Mathematics Teaching	**11**	**.80**
Influence on Teaching - Standards and Testing (subscale)	4	.85
Professional Interaction	**5**	**.80**
Professional Interaction with M^2 Teacher Leaders	**8**	**.77**

Preparedness for Mathematics Instruction

	N	Mean 2004	Mean 2005	Mean 2006	Difference 2004-05	Difference 2005-06	Difference 2004-06
Overall Preparedness for Teaching Mathematics	**28**	**2.47**	**2.87**	**3.21**	**.40*** **	**.34*** **	**.74*** **
Preparedness to Teach Diverse Populations	**28**	**2.21**	**2.51**	**2.72**	**.30***	**.21**	**.51*** **
Use action research.	28	1.68	1.86	3.43	.18	1.57***	1.75***
Use a variety of assessment strategies.	28	2.21	2.79	3.43	.57**	.64***	1.21***
Use student assessment results.	28	2.57	2.96	3.43	.39	.46*	.86***
Teach mathematics with technology tools.	28	1.96	2.29	2.82	.32	.54	.86***
Teach mathematics with manipulative materials.	28	2.26	2.70	3.11	.44*	.41*	.85***
Teach problem-solving strategies.	28	2.50	3.14	3.29	.64***	.14	.79***
Select/adapt instructional materials.	28	2.79	3.36	3.50	.57***	.14	.71**
Sequence mathematics instruction.	28	2.70	3.26	3.41	.56**	.15	.70***
Encourage participation of minorities.	28	2.64	3.00	3.32	.36*	.32	.68***
Provide a challenging curriculum for all students.	28	2.89	3.36	3.54	.46**	.18	.64***
Provide instruction that meets challenging standards.	27	2.96	3.44	3.56	.48**	.11	.59***
Teach students with diverse abilities.	27	2.59	2.96	3.15	.37	.19	.56**
Teach students with learning disabilities.	28	2.29	2.54	2.82	.25	.29	.54**
Teach students with limited English proficiency.	28	1.68	1.93	2.21	.25*	.29	.54**
Connect mathematics and other subject areas.	28	2.71	3.14	3.21	.43*	.07	.50**
Encourage participation of females.	28	3.25	3.46	3.68	.21	.21	.43**
Teach students with a variety of cultural backgrounds.	28	2.32	2.68	2.71	.36	.04	.39*

Note: Responses were rated on a 4-point scale where 1 = Not Well Prepared, 2 = Somewhat Prepared, 3 = Well Prepared, 4 = Very Well Prepared. Bold text indicates composite variables. The 2004-2006 difference may not equal the sum of the annual differences due to rounding. *p < .05, **p < .01, ***p < .001.

RMC RESEARCH

Confidence in Mathematics Instruction and Teacher Support

	N	Mean 2004	Mean 2005	Mean 2006	Difference 2004-2005	Difference 2005-2006	Difference 2004-2006
Overall Confidence in Mathematics and Teacher Support	**29**	**2.57**	**2.94**	**3.35**	**.38****	**.40*****	**.78*****
Confidence in Ability to Support Colleagues	**29**	**2.39**	**2.86**	**3.30**	**.47****	**.45*****	**.92*****
Confidence in Mathematical Knowledge	**28**	**2.69**	**3.04**	**3.42**	**.35*****	**.38*****	**.73*****
Confidence in Leadership Ability	**29**	**2.85**	**3.03**	**3.36**	**.19**	**.33***	**.52*****
Knowledge about educational issues related to mathematics	29	2.14	2.66	3.31	.52**	.66***	1.17***
Ability to provide multiple types of support to colleagues	29	2.14	2.55	3.31	.41	.76***	1.17***
Ability to coach or mentor new teachers	29	2.66	3.31	3.59	.66**	.28	.93***
Ability to help colleagues improve mathematics knowledge and skills	28	2.50	3.18	3.36	.68***	.18	.86***
Ability to write mathematics curriculum	29	2.35	2.72	3.17	.38	.45*	.83***
Ability to coach or mentor experienced teachers	29	2.28	2.55	3.10	.28	.55***	.83***
Ability to act as a leader among other teachers	29	2.86	3.07	3.45	.21	.38*	.59***
Knowledge beyond what you teach	28	2.54	2.86	3.07	.32	.21	.54***
Other teachers see you as a leader	29	2.83	3.00	3.28	.17	.28	.45**
Knowledge related to mathematics you teach	27	3.48	3.63	3.89	.15	.26	.41**

Note: Responses were rated on a 4-point scale where 1 = Not Confident at All, 2 = Somewhat Confident, 3 = Moderately Confident, 4 = Very Confident. Bold text indicates scaled items. The 2004-2006 difference may not equal the sum of the annual differences due to rounding. *p < .05, **p < .01, ***p < .001.

RMC RESEARCH

Summary of Findings Across 4 Cohorts

Teachers . . .

- Were more prepared and confident to teach mathematics and provide support to other teachers;
- Deemphasized the need for basic mathematics skills, use of algorithms, and repeated practice;
- Increased instructional emphasis on NCTM process standards such as communication, representation, and connections;
- Increased use of assessment activities, including those that emphasized assessment and justification and those involving demonstration and performance; and
- Increased their professional interaction among colleagues, including discussions about how to teach, collaborating to prepare instructional materials, and observing colleagues' teaching.

M^2 Activities . . .

- Enhanced faculty knowledge and interests regarding K-12 schools and teachers;
- Were consistently aligned with mathematics content and process standards and received high ratings from participants; and
- Aligned well with multiple indicators of sustainability.

RMC RESEARCH

Summary of Findings: Teacher Outcomes

	Two-Year Gain			
	Cohort 1	Cohort 2	Cohort 3	Cohort 4
Overall Preparedness for Teaching Mathematics	+	+	+	+
Preparedness to Teacher Diverse Students (subscale)	+	+	+	+
Overall Confidence in Mathematics and Teacher Support	+	+	+	+
Confidence in Mathematical Knowledge (subscale)	+	+	+	+
Confidence in Ability to Support Colleagues (subscale)	+	+	+	+
Confidence in Leadership Ability (subscale)	+	+	+	+
Deemphasis on Need for Basic Mathematics Skills, Memorization, Use of Algorithms, and Repeated Practice (item-level analysis)	+	+	+	+
Emphasis on NCTM Process Standards	+	+	+	+
Increased use of Instructional Activities such as Working in Small Groups, Working on Problems that Take Over 30 Minutes to Solve, and Involve Explanations of Mathematical Reasoning (item-level analysis)	+	+	+	+
Overall Use of Assessment in Mathematics	+	+	+	+
Use of Assessment – Analysis and Justification (subscale)	+	+	+	+
Professional Interaction	+	+	+	n.s.

Note: + = statistically significant positive effect at the $p < .05$ level; n.s. = not significant.

RMC RESEARCH

Student Mathematics Achievement Outcomes

Student Achievement Data

- Provided by LPS for all Grade 5-8 students during 2004-05, 2005-06, and 2006-07 school years
- Spring 2004, 2005, 2006, 2007 math scores
  - District designed CRT (grades 4 and 8); total math added for grades 5-7 in 2006
  - Metropolitan Achievement Test (Grades 5-7)

Student Achievement Data

- Scale scores on MAT: *concepts and problem solving, procedures, total math*
- Raw scores on CRT: *algebra, computation, data analysis, geometry, measurement, and numeration*

Student Achievement Data

- Student identifiers allow linkage over time (analyses control for prior achievement)
- Teacher identifiers (grades 6-8) allow linkage to teacher survey data
- Student demographic information (e.g., gender, race/ethnicity, participation in special programming)

LPS Student Achievement Measures

	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Total Math (MAT scale score)		X	X[a]	X	
Math Procedures		X	X[a]	X	
Math Concepts and Problem Solving		X	X[a]	X	
Total Math (District CRT point total)	X	X[b]	X[b]	X[b]	X
Algebra	X		X[b]		X
Computation	X		X[b]		X
Data Analysis	X		X[b]		X
Geometry and Measurement	X		X[b]		X
Numeration	X		X[b]		X

[a]2004, 2005, and 2006 only; [b]2006 and 2007 only

LPS Student Samples (2006-07)

	Teachers	All Students	M^2 Students	
			n	Percent
Grade 6	1 M^2 Cohort 1 1 M^2 Cohort 2 3 M^2 Cohort 3 99 Comparison	2,188	172	7.9
Grade 7	3 M^2 Cohort 1 3 M^2 Cohort 2 0 M^2 Cohort 3 35 Comparison	2,228	383	17.2
Grade 8	7 M^2 Cohort 1 4 M^2 Cohort 2 1 M^2 Cohort 3 34 Comparison	2,257	921	40.8

LPS Student Demographics, Grade 8

	2006-2007			
	Students of M² Participants (*N* = 921)		Students of Comparison Teachers (*N* = 1,368)	
	n	Percent	*n*	Percent
Gender				
Female	463	50.3	668	48.8
Male	458	49.7	700	51.2
Race/Ethnicity				
White	774	84.0	1,078	78.8
African American	51	5.5	130	9.5
Hispanic	41	4.5	90	6.6
Asian	46	5.0	52	3.8
Other	9	1.0	18	1.3
LPS Program Participation				
Gifted and Talented	275	29.9	207	15.1
Special Education	118	12.8	243	17.8
English Language Learner	27	2.9	59	4.3
Course Enrollment				
Below Grade Level	113	12.3	260	19.0
On Grade Level	591	64.2	950	69.4
Above Grade Level	329	35.7	222	16.2

Influence of M² Participation on 2007 Achievement, Grade 8

	Data Analysis (District CRT)		Geometry and Measurement (District CRT)		Numeration (District CRT)	
	B	*SE*	*B*	*SE*	*B*	*SE*
Intercept	5.88***	0.20	5.26***	0.42	2.63***	0.21
2006 Math Achievement (District CRT Total Math point total)	0.07***	0.00	0.20***	0.01	0.10***	0.00
Student of M2 Teacher (1 = in M2 classroom during 2006-2007 school year)	-0.04	0.09	0.48*	0.19	0.13	0.10
Gender (1 = male)	0.03	0.07	0.20	0.14	0.18*	0.07
African American	-0.70***	0.14	-0.44	0.29	-0.03	0.15
Hispanic	-0.28~	0.16	0.29	0.33	-0.32~	0.17
Asian	-0.07	0.18	0.21	0.36	-0.08	0.18
Gifted and Talented	0.14	0.10	1.39***	0.19	0.18~	0.10
Special Education	-0.53***	0.12	-0.46~	0.24	-0.18	0.12
English Language Learner	-0.26	0.22	0.84~	0.45	0.14	0.23
R 2	.42		.57		.53	
Number of Observations	1,975		1,975		1,975	

~p<.10, *p < .05, **p < .01, ***p < .001.

Summary of Effects: LPS Student Achievement (2004-2005)

	Effect on 2005 Mathematics Achievement Associated With M² Participation		
	Grade 6	Grade 7	Grade 8
Total Math (MAT NCE score)	n.s.	n.s.	Test Not Administered
Math Procedures	n.s.	n.s.	Test Not Administered
Math Concepts and Problem Solving	n.s.	n.s.	Test Not Administered
Total Math (District CRT point total)	Test Not Administered	Test Not Administered	Positive (.20)
Algebra	Test Not Administered	Test Not Administered	Positive (.20)
Computation	Test Not Administered	Test Not Administered	Positive (.14)
Data Analysis	Test Not Administered	Test Not Administered	n.s.
Geometry and Measurement	Test Not Administered	Test Not Administered	Positive (.18)
Numeration	Test Not Administered	Test Not Administered	Positive (.13)

Note: n.s. = not significant at the $p < .05$ level. Effect size (standardized mean difference) is indicated in parentheses.

Summary of Effects: LPS Student Achievement (2005-2006)

	Effect on 2006 Mathematics Achievement Associated With M² Participation		
	Grade 6	Grade 7	Grade 8
Total Math (MAT NCE score)	n.s.	n.s.	Test Not Administered
Math Procedures	n.s.	n.s.	Test Not Administered
Math Concepts and Problem Solving	n.s.	n.s.	Test Not Administered
Total Math (District CRT point total)	n.s.	Positive (.14)	n.s.
Algebra	n.s.	Test Not Administered	n.s.
Computation	n.s.	Test Not Administered	n.s.
Data Analysis	n.s.	Test Not Administered	n.s.
Geometry and Measurement	n.s.	Test Not Administered	n.s.
Numeration	n.s.	Test Not Administered	n.s.

Note: n.s. = not significant at the $p < .05$ level. Effect size (standardized mean difference) is indicated in parentheses.

Summary of Effects: LPS Student Achievement (2006-2007)

	Effect on 2007 Mathematics Achievement Associated With M² Participation		
	Grade 6	Grade 7	Grade 8
Total Math (MAT NCE score)	Test not administered	n.s.[a]	Test not administered
Math Procedures	Test not administered	Positive (.13)	Test not administered
Math Concepts and Problem Solving	Test not administered	n.s.	Test not administered
Total Math (District CRT point total)	Negative (-.09)	Positive (.15)	n.s.
Algebra	n.s.	Test not administered	n.s.
Computation	Negative (-.16)	Test not administered	n.s.
Data Analysis	n.s.	Test not administered	n.s.
Geometry and Measurement	n.s.	Test not administered	Positive (.10)
Numeration	Negative (-.21)	Test not administered	n.s.

Note: n.s. = not significant at the $p < .05$ level. Effect size (standardized mean difference) is indicated in parentheses.

Challenges Linking Teacher Prof. Devel. And Practice to Student Achievement

- Review of over 1,300 studies that examined the effect of teacher professional development on student achievement (Yoon et al., 2007) found that only nine met national (WWC) standards for rigorous evidence.
- Recent study of "reform oriented instruction" (Le et al, 2009)
  - Weak relationship to student achievement
  - Issues
    - Alignment with state/district tests
    - Open-ended items measuring problem solving skills
    - Use of test subscores

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (Issues & Answers Report, REL 2007–No. 033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest.

Le, V., Lockwood. J.R., Stecher, B.M., Hamilton, L.S., Martinez, J.F. (2009). A longitudinal investigation of the relationship between teachers' self-reports of reform-oriented instruction and mathematics and science achievement. *Educational Evaluation and Policy Analysis*, 31(3).

Limitations

- Self-selection; small sample of teachers
- No baseline data for comparison teachers
- Generalizability for some grade levels that had relatively small numbers of M^2 participants
- Good student achievement data available for only one district
- Low statistical power for hierarchical analysis
- Lack of data to control for school and student level socioeconomic status

LPS is "Best Case" for Achievement Data!

Collecting achievement data from the non-LPS districts continues to be a challenge. First effort to collect data from the 11 non-LPS districts represented in Cohort 1 resulted in the following:

- Six districts returned achievement data (TerraNova, NWEA, SAT); scale score, percentile ranks, and NCE scores.
- In four districts, M^2 participants were the only mathematics teacher(s) at a particular grade level. In the 2 districts with comparison students, there were fewer than 20 students in either the M^2 or comparison group.
- Only one district provided individual-level student data that could be linked across years for students of M^2 participants and nonparticipants.
- None provided demographic or other student data that could be used to control for other possible influences on achievement.
- The limitations of the non-LPS data (i.e., inconsistent outcome measures, inadequate comparison groups, and limited ability to link student data over time) allow only very weak conclusions about the impact of M^2.

Next Steps

As additional longitudinal data are collected from M^2 participants, comparison teachers, and their students, analyses will better allow for conclusions about impact.

- Continued efforts to collect and aggregate data from rural districts
- More sophisticated assignment of students to teachers
- Options for aggregating teacher data across cohorts and student data across years
- Multiyear gains
- "Concentration effects"

Feedback or Questions

Stephen Meyer
meyer@rmcdenver.com

John Sutton
sutton@rmcdenver.com