

**Project CoMPASS
Evaluation Report**

July 2016



LeCroy & Milligan
ASSOCIATES, INC.

Project CoMPASS Evaluation Report July 2016

Submitted to:

Flowing Wells Unified School District
1556 W. Prince Rd.
Tucson, AZ 85705
Ph: (520) 696-8800
Fax: (520) 690-2400
www.floatingwellsschools.org

Submitted by:

LeCroy & Milligan Associates, Inc.
2002 N. Forbes Blvd. Suite 108
Tucson, AZ 85745
Ph: (520) 326-5154
Fax: (520) 326-5155
www.lecroymilligan.com



Acknowledgments:

The evaluation team for this project would like to thank Michael Perkins, Project Director, and Kevin Stoltzfus, Project Co-PI and Assistant Superintendent of Flowing Wells Unified School District, for their extensive efforts in implementing Project CoMPASS and guidance with the evaluation. The evaluation team also thanks the project's instructional team, Jason Aubrey and Rebecca McGraw from the University of Arizona, for their participation in collecting evaluation data. The evaluation team includes Steven Wind, Ph.D., Jenny McCullough Cosgrove, MSW, Sonia Cota-Robles, Ph.D., JD., Michele Schmidt, MPA, and Eloina Cardenas.

About LeCroy & Milligan Associates:

Founded in 1991, LeCroy & Milligan Associates, Inc. is a consulting firm specializing in social services and education program evaluation and training that is comprehensive, research-driven and useful. Our goal is to provide effective program evaluation and training that enables stakeholders to document outcomes, provide accountability, and engage in continuous program improvement. With central offices located in Tucson, Arizona, LeCroy & Milligan Associates has worked at the local, state and national level with a broad spectrum of social services, criminal justice, education and behavioral health programs.

Suggested Citation:

LeCroy & Milligan Associates, Inc. (2016). *Project CoMPASS Evaluation Report*. Tucson, AZ: Author.



Report Contents

Executive Summary	3
Introduction	7
Teacher Recruitment and Retention	9
Evaluation Methodology.....	11
Group Equivalence.....	11
Instruments and Measures.....	11
Data Collection	13
Data Analysis.....	14
Report Structure	14
Teacher Findings	15
MMTsm	15
RTOP	16
Use of District Benchmark Data	17
Development of Teaching Resources	17
Application of New Pedagogical Strategies	18
Student Findings	19
District Benchmarks.....	19
Teacher Perceptions of Student Reaction to Use of New Instructional Strategies.....	21
Appendix A.....	23

List of Exhibits

Exhibit 1. Project CoMPASS Goals and Objective	8
Exhibit 2. Number of Experimental and Control Group Teachers by District and School Affiliation	10
Exhibit 3. Lesson Developed by Participants.....	17
Exhibit 4. FUSD Participants' Student Benchmark Change	20
Exhibit 5. MUSD Participants' Student Benchmark Change.....	20
Exhibit 6. SUSD Participants' Student Benchmark Change	20



Executive Summary

Flowing Wells Unified School District (FUSD), in partnership with Marana Unified School District (MUSD) and Sahuarita Unified School District (SUSD), was awarded a Mathematics and Science Partnership (MSP) Grant by the Arizona Department of Education to provide professional development to mathematics teachers in grades 7-12. The consortium named its professional development effort Project CoMPASS (Conceptual Mathematics Partnership for Arizona's Secondary Students). The three school districts in the consortium had identified a need for high quality mathematics-focused professional development based on student achievement results. The overarching focus of the professional development was on pedagogical practices to develop students' conceptual understanding and unpack misconceptions, which aligns with the emphasis in the Arizona College and Career Ready Standards on problem solving and mathematical practices.

FUSD, acting as the lead agency in the project, coordinated the recruitment of teachers. The project's co-directors held meetings with assistant superintendents from all three districts, at which they provided a flyer with information about Project CoMPASS and requested that the assistant superintendents meet with the teachers in the math departments of their districts' middle and high schools to encourage participation. FUSD followed up with the contact made by the assistant superintendents with recruitment e-mails to all middle and high school math teachers. When a sufficient number of control teachers was not recruited from the three partnering districts in this manner, the project sought out the participation of another district with similar socio economic demographics. The project was able to identify and obtain the cooperation of Amphitheater Unified School District, using the same recruitment methods as had been used at the three original partnering districts.

Using these strategies, the project was able to recruit a convenience sample of 33 teachers as participants (referred to in this report as "experimental teachers") and 31 teachers to be in a control group. Twenty-nine of the experimental teachers completed at least 90% of the professional development, for an 88% retention rate. On average, the experimental teachers completed 100.3 of the 104 hours of professional development provided. Thirty of the thirty-one control group teachers completed the project, for a 97% retention rate for the control group.

The experimental teachers taught a total of 3,221 students, 1,920 middle school students and 1,301 high school students.

The following narrative highlights the key findings of the evaluation for each of the project's goals and objectives.



Goal 1: Develop a cadre of highly competent teachers of grade 7-12 mathematics.

Objective 1.1: By April 29, 2016, all participating teachers will demonstrate at least 10% growth in their mathematical content knowledge, as measured by an approved teacher content assessment (such as the Mathematical Meanings for Teaching secondary mathematics instrument, or MMTsm).

Finding: 52% of participating teachers who completed the professional development achieved a 10% growth in their mathematical content knowledge, as evidenced by a change in their MMTsm score, falling short of the target for all participants to achieve such growth.

Objective 1.2: By April 29, 2016, all participating teachers will demonstrate at least 10% growth in their propositional pedagogic knowledge, as measured by ratings on the “Content - Procedural Pedagogic Knowledge” subscale on the Reformed Teaching Observation Protocol (RTOP).

Finding: 58% of participating teachers achieved a 10% or greater increase in their propositional pedagogic knowledge, as measured by ratings on the Content - Procedural Pedagogic Knowledge subscale, falling short of the target for all participants to achieve such growth.

Goal 2: Improve student achievement in mathematics in participating teachers’ classrooms.

Objective 2.1: By April 29, 2016 (pending availability of scores), students in participating teachers’ classrooms will demonstrate a 10% increase in the percentage who score proficient (Meets or Exceeds) on the 2016 AzMERIT state assessment compared to the 2015 AzMERIT state assessment. Each student’s 2016 score will be compared against his or her 2015 score.

Finding: AzMERIT data were not obtained, precluding reporting on whether this objective was met. Following the completion of the project, the project’s co-directors and ADE staff overseeing the MSP program collaboratively decided that AzMERIT data would not serve as an appropriate measure of student outcomes because there were questions as to whether the data are valid and reliable. Specifically, it was the first time students took this assessment, it was quite different from AIMS, and it was administered in different ways (paper and by computer). Additionally, student motivation for the AzMERIT likely was low given that the high-stakes graduation requirement associated with the former AIMS test had been removed for AzMERIT. Finally, it would have been difficult to obtain AzMERIT scores from partner districts in a timely fashion because scores were to be released during the summer break.



Objective 2.2: By April 29, 2016, students in participating teachers’ classrooms will demonstrate a 10% increase in the percentage who score proficient (Meets or Exceeds) on 2016 district/school benchmark assessments compared to 2015 benchmark assessments.

Finding: Only aggregate (i.e., average) benchmark data were available for participants’ students, making it impossible to assess change in the percent of students who scored proficient on the benchmarks. Moreover, all districts did not use the same benchmarks and scores for the same subject for 2015 and 2016, and data were only available for 20 experimental teachers. The results of the limited analysis possible do not support a finding that the professional development contributed to an increase in student benchmark scores.

Goal 3: Improve teachers’ ability to frequently monitor the learning of all students and adjust instruction to better meet their needs.

Objective 3.1: By April 29, 2016, all participating teachers will demonstrate at least 10% growth in their ability to implement formative, in-class assessments and to provide appropriate instruction as measured by ratings on the “Classroom Culture - Communicative Interactions” subscale on the RTOP.

Finding: 58% of participating teachers achieved a 10% or greater increase in their propositional pedagogic knowledge, as measured by ratings on the Classroom Culture - Communicative Interactions subscale, falling short of the target for all to achieve such growth.

Objective 3.2: By April 29, 2016, all participating teachers will use district benchmark assessment results to adjust curriculum and instruction.

Finding: FWUSD’s middle school and high school math departments, which included all Project CoMPASS participants, met following each round of benchmarks to analyze results and determine next steps. However, teachers’ made limited adjustments to their curricula and instruction due to a sense of urgency to move forward to completion of those curricula. Data are not available on how SUSD and MUSD participants made use of district benchmark data.



Goal 4: Develop professional development resources to be shared with other educators through an online environment.

Objective 4.1: By April 29, 2016, all participating teachers will have developed supporting activities, lessons, and formative assessments congruent to the mathematical content addressed throughout the professional development; each participating grade level or course will have developed at least 10 lessons and supporting assessments and posted all resources in a shared online environment.

Finding: Participants developed 50 lessons with supporting activities and formative assessments which an instructor posted on the projects Google Group site. The teachers developed lessons for Grade 7, Grade 8, Algebra 1, Geometry, and Algebra 2/Pre-calculus.



Introduction

Flowing Wells Unified School District (FUSD) partnered with Marana Unified School District (MUSD) and Sahuarita Unified School District (SUSD) to form a consortium to implement a Mathematics and Science Partnership (MSP) Grant professional development project for mathematics teachers in grades 7-12. The consortium named this work Project CoMPASS (Conceptual Mathematics Partnership for Arizona's Secondary Students). Each of the three school districts in the consortium had identified a need for high quality mathematics-focused professional development based on student achievement results. The districts had experienced high failure rates by students in high school level algebra courses and only one of the eleven participating schools achieved the 2014 Annual Measurable Objectives (AMOs) in mathematics. Results of first semester benchmarks for 2014-2015 conducted with the ATI Galileo platform benchmarks also demonstrated inconsistent achievement of the grade-level AMOs in some participating schools. Analysis of the 2014 AIMS results at the Strand and Concept level provided additional insights into students' areas of weakness and suggested potential areas for focus during teacher professional development.

During the preliminary planning phase of the project, the districts surveyed middle and high school teachers to gain additional information about professional development needs. However, the overarching focus of the professional development was on pedagogical practices to develop students' conceptual understanding and to unpack misconceptions, which align with the emphasis in the Arizona College and Career Ready Standards on problem solving and mathematical practices.

Dr. Kevin Stoltzfus, Assistant Superintendent of FUSD, and Michael Perkins, the FUSD math specialist, served as co-directors of the project. The University of Arizona served as the Institute of Higher Learning for the project, with Dr. Cody Patterson providing mathematics content instruction and Dr. Rebecca McGraw providing mathematics pedagogy instruction. Dr. Jason Aubry joined the project when Dr. Patterson took a position at another university. Project CoMPASS had four goals, each having one to two objectives, which are shown in Exhibit 1.



Exhibit 1. Project CoMPASS Goals and Objective

Goal	Objective
<p>Goal 1: Develop a cadre of highly competent teachers of grade 7-12 mathematics.</p>	<p>Objective 1.1: By April 29, 2016, all participating teachers will demonstrate at least 10% growth in their mathematical content knowledge, as measured by an approved teacher content assessment (such as the Mathematical Meanings for Teaching secondary mathematics instrument or MMTsm).</p>
	<p>Objective 1.2: By April 29, 2016, all participating teachers will demonstrate at least 10% growth in their propositional pedagogic knowledge, as measured by ratings on the “Content - Procedural Pedagogic Knowledge” subscale on the Reformed Teaching Observation Protocol (RTOP).</p>
<p>Goal 2: Improve student achievement in mathematics in participating teachers’ classrooms.</p>	<p>Objective 2.1: By April 29, 2016 (pending availability of scores), students in participating teachers’ classrooms will demonstrate a 10% increase in the percentage who score proficient (Meets or Exceeds) on the 2016 AzMERIT state assessment compared to the 2015 AzMERIT state assessment. Each student’s 2016 score will be compared against his or her 2015 score.</p>
	<p>Objective 2.2: By April 29, 2016, students in participating teachers’ classrooms will demonstrate a 10% increase in the percentage who score proficient (Meets or Exceeds) on 2016 district/school benchmark assessments compared to 2015 benchmark assessments.</p>
<p>Goal 3: Improve teachers’ ability to frequently monitor the learning of all students and adjust instruction to better meet their needs.</p>	<p>Objective 3.1: By April 29, 2016, all participating teachers will demonstrate at least 10% growth in their ability to implement formative, in-class assessments and to provide appropriate instruction as measured by ratings on the “Classroom Culture - Communicative Interactions” subscale on the Reformed Teaching Observation Protocol (RTOP).</p>
	<p>Objective 3.2: By April 29, 2016, all participating teachers will use district benchmark assessment results to adjust curriculum and instruction.</p>
<p>Goal 4: Develop professional development resources to be shared with other educators through an online environment.</p>	<p>Objective 4.1: By April 29, 2016, all participating teachers will have developed supporting activities, lessons, and formative assessments congruent to the mathematical content addressed throughout the professional development; each participating grade level or course will have developed at least 10 lessons and supporting assessments and posted all resources in a shared online environment.</p>

Following the completion of the project, the Project CoMPASS co-directors and Arizona Department of Education (ADE) staff overseeing the MSP program collaboratively decided that AzMERIT data would not serve as an appropriate measure of student outcomes (see Objective 2.1). This decision was made for several reasons. First, it was unclear whether the



2015 AzMERIT was valid and reliable. This school year was the first time students had taken the test, which was fundamentally different from the previous state assessment (AIMS). Students at different schools took different versions of the AzMERIT: some students completed the test on paper, while others took a computer-based version of the test. Additionally, student motivation for the AzMERIT likely was low given that the high-stakes graduation requirement associated with the former AIMS test had been removed for AzMERIT. Finally, the late availability of the 2016 AzMERIT scores also contributed to the decision to use other student outcome measures. The AzMERIT scores would not be available until June 6, 2016, making it difficult to obtain scores from partner districts in a timely fashion during the summer break.

Teacher Recruitment and Retention

FUSD, acting as the lead agency in the project, coordinated the recruitment of teachers. The project's co-directors held meetings with assistant superintendents from all three districts. At these meetings, assistant superintendents received an informational flier about Project CoMPASS and were requested to meet with teachers in the math departments of their districts' middle and high schools to encourage participation in the project. FUSD followed up with the contact made by the assistant superintendents with recruitment e-mails to all middle and high school math teachers.

The project utilized a quasi-experimental matched treatment and control group study design. Teachers in the experimental group were offered a stipend upon completion of professional development; teachers in the control group received a different level of stipend based on having a completed pre- and post- content assessment and a classroom observation. Recruitment of a sufficient number of control teachers within the three districts proved to be difficult. The project overcame this challenge by recruiting teachers for the control group from the Amphitheater Unified School District (AUSD), a district with similar socioeconomic, racial, and ethnic demographics as the treatment group districts. The original convenience sample consisted of 33 experimental group teachers and 31 control group teachers. Exhibit 2 shows the number of participant and control group teachers by district and school.



Exhibit 2. Number of Experimental and Control Group Teachers by District and School Affiliation

District	School	Number of Teachers	
		Experimental Group	Control Group
Flowing Wells Unified School District	Flowing Wells High School	7	5
	Flowing Wells Junior High School	6	3
	Sentinel Peak High School	0	1
Marana Unified School District	Marana High School	3	1
	Marana Middle School	2	5
	Marana Career and Technical High (MCAT) High School	1	0
	Mountain View High School	3	2
	Tortolita Middle School	6	0
Sahuarita Unified School District	Anza Trail School	2	0
	Sahuarita Middle School	3	1
Amphitheater Unified School District	Amphitheater High School	0	7
	Amphitheater Middle School	0	4
	La Cima Middle School	0	2
Total		33	31

Project completion for the experimental group teachers was defined as having completed at least 90% or 93.6 hours of the 104 hours of professional development provided. Based on this definition, 29¹ of the 33 experimental teachers recruited completed the project, for an 88% completion rate and 12% attrition rate. On average, the experimental group completed 100.3 hours of professional development.

Completion for control group teachers was defined as having completed both a pre- and post- content assessment. The control group had a 97% retention rate and 3% attrition rate². One control group teacher left later in the project due to a scheduling conflict.

¹ While four experimental group teachers did not complete at least 90% of the professional development, two came close by completing 88% (92.0 hours).

² Two control teachers dropped out very early in the project, before ID numbers were even assigned to teachers, and were therefore not counted as recruited or included in completion or attrition rate computations.



The overall attrition rate for the project is based on the combined number of teachers that left both groups. Five teachers in total dropped out of the project, for an overall attrition rate of 8%.

Evaluation Methodology

Group Equivalence

The project collected basic demographic data from recruited treatment and control group teachers. The evaluation team assessed the baseline equivalency of these two groups, comparing them by age, gender, grades teaching, years of teaching experience, and whether they held a math-related degree. Chi square analysis found no significant differences between the groups, indicating the two groups were well matched on these variables.

Instruments and Measures

Utilizing a pre/post quasi-experimental design, experimental and control group teachers were compared by: measures of content and propositional pedagogic knowledge (Goal 1); student achievement (Goal 2); teaching skills (Goal 3); and developed teaching resources (Goal 4). Three instruments described below were used to gather data on the project's impact on teachers. Additionally, the data used to measure student achievement is also described below.

Content Knowledge

The Mathematical Meanings for Teaching Secondary Mathematics (MMTsm) is a 44-item diagnostic instrument that assesses the mathematical meanings with which teachers operate. The MMTsm organizes diagnostic items according to nine constructs: Covariation; Function (General Concept); Functions as Models; Function Notation; Frames of Reference; Magnitude; Proportionality; Rate of Change; and Structure.

Teaching Skill

The Reformed Teacher Observation Protocol (RTOP) consists of five 5-item subscales that measure the degree to which certain student-centered engaged learning practices are present in a classroom. The five subscales are: Lesson Design and Implementation; Propositional Knowledge; Procedural Knowledge; Communicative Interactions; Student/Teacher Relationships; and Lesson Design and Implementation. Each item is scored from 0 to 4, for a maximum total score of 100 for 25 items. The project's objectives related to two of the subscales: Procedural Knowledge and Communicative Interactions.



The Procedural Knowledge subscale of the RTOP consists of five items:

- Students used a variety of means (models, drawings, graphs, symbols, concrete materials, manipulatives, etc.) to represent phenomena;
- Students made predictions, estimations and/or hypotheses, and devised means for testing them;
- Students were actively engaged in thought-provoking activity that often involved the critical assessment of procedures;
- Students were reflective about their learning; and
- Intellectual rigor, constructive criticism, and the challenging of ideas were valued.

The Communicative Interactions subscale of the RTOP also consists of five items:

- Students were involved in the communication of their ideas to others using a variety of means and media;
- The teacher's questions triggered divergent modes of thinking;
- There was a high proportion of student talk and a significant amount of it occurred between and among students;
- Student questions and comments often determined the focus and direction of classroom discourse; and
- There was a climate of respect for what others had to say.

Participant Survey

The project developed a survey that participants completed during the last day of professional development (this survey is available in Appendix A of this report). Areas addressed by 11 open-ended items include: use and sharing of and supported needed for implementing new teaching strategies learned; response of students to use of new teaching strategies; most useful and challenging aspects of the professional development; and suggestions for improving the professional development. Four multiple choice items examined change in teacher confidence in teaching mathematics, and satisfaction with the professional development's pacing and printed materials.

Additional teacher data for the evaluation was obtained through review of lessons developed by participants that were posted on a Google Group site started by the project.

Student Achievement

Project CoMPASS originally planned to use AzMERIT data to assess the project's impact on student achievement. However, following the completion of the project, the project's co-directors and ADE staff overseeing the MSP program collaboratively decided that AzMERIT data would not serve as an appropriate measure of student outcomes. This decision was made for several reasons. First, it was unclear whether the 2015 AzMERIT



was valid and reliable because it was the first time students had taken the test, which was fundamentally different from the previous state assessment (AIMS). Moreover, students at different schools took different versions of the AzMERIT, some completing a traditional paper test and others taking a computer-based version. Additionally, student motivation for the AzMERIT likely was low given that the high-stakes graduation requirement associated with the former AIMS test had been removed for AzMERIT. The fact that the 2016 scores would not be available until during the summer break, making it difficult to obtain scores from some participating districts, further supported the decision.

Instead, the project used district benchmarks to assess impact on student achievement, with each participating district having its own form of benchmark and scoring. FWUSD and SUSD used benchmarks developed in coordination with Assessment Technology Incorporated (ATI) (i.e., ATI's Galileo assessment system). The benchmarks mostly consisted of items from ATI's item bank, with a few modifications to meet each district's needs. MUSD used the second semester final exam grade as its benchmark for high school students but was unable to provide benchmark data for its junior high school students.

Data Collection

The project's IHE instructors administered the MMTsm to both experimental and control teachers. The pre-MMTsm was administered to experimental teachers between May 6 and May 12, 2015 and to control teachers between September 1 and 14, 2015. The later period for administration to control teachers was necessary to allow for adequate recruitment for that group. All but one experimental group teachers and all but two control group teachers completed the post-MMTsm between March 3 and 8, 2016. Teachers who were not available for the early March testing dates completed the post- MMTsm between April 6 and 13, 2016.

Three members of the evaluation team participated in an ADE-sponsored RTOP training prior to launch of the project and were responsible for conducting all RTOP observations. To establish interrater reliability, soon after the training the observers as a group viewed on-line classroom videos recommended by the RTOP trainer for further RTOP practice. They compared their scoring and, through discussion, established a consensus on how items that had been scored differently should be scored to be in agreement with a scoring rubric provided by ADE. To further strengthen interrater reliability, the three observers conducted the first three observations as a group, again comparing scoring and reaching consensus on items that were scored differently.

The evaluation team conducted all pre- RTOP observations for the experimental teachers between 4/23/15 and 5/12/15. Observations for all but one of the control teachers occurred between 8/24/15 and 9/8/15, to allow for adequate recruitment for that group.



One control teacher was on maternity leave for the first months of the school year and could not be observed until 11/10/15. Post- RTOP observations for experimental and control teachers occurred between 4/11/16 and 5/5/16. Due to the previously described attrition and a control teacher only teaching a computer lab during the second semester of the school year, matched pre- and post- RTOP data were available for 31 experimental group teachers and 28 control group teachers.

One of project's co-directors was responsible for obtaining student benchmark data for the current and previous academic year for students taught by experimental teachers from each of the collaborating districts.

Data Analysis

The project's IHE instructors administered the MMTsm to all experimental and control teachers and generated raw scores for all sections of the assessment. However, full analysis of MMTsm scores is a complex process. In past years, a statistician from Arizona State University, where the MMTsm was developed, performed the analysis for all MSP projects that used this instrument. The co-principal investigators decided that it would be prudent to have the same individual do the analysis of MMTsm scores for Project CoMPASS.

The evaluation team conducted the analysis of RTOP scores. All RTOP data was imported into SPSS 22 and change scores were computed to determine whether the project objective of all experimental teachers showing a 10% increase in content knowledge was met.

The evaluation team received student benchmark data from the project's co-director in an Excel worksheet. Given the small amount of data available and the limited analysis possible, all data cleaning and data analysis were performed in Excel, rather than SPSS.

Report Structure

Findings from this project for teachers and students are presented in separate sections. The 35-page report containing the full analysis of MMTsm results that was completed by Neil Hatfield, a statistician at the Department of Mathematical and Statistical Sciences of Arizona State University, has been submitted as an attachment to the project's APR.



Teacher Findings

MMTsm

The MMTsm is a 44-item diagnostic instrument that assesses the mathematical meanings with which teachers operate. To achieve the most robust analysis possible of teachers' performance on the MMTsm, the analyst created u^* scores based on the data. A u^* score is a ranking of each teachers performance on the MMTsm as compared with that of every other teacher. Analysis using standard ranking compares each participant's performance against those of other participants using only one variable. However, for the analysis of Project CoMPASS MMTsm data, items were organized into nine constructs: co-variation, function, function as model, function notation, frame of reference, magnitude, proportionality, rate of change, and structure. In this way, the analysis produced nine separate u^* scores, one for each construct, allowing a more robust comparison of teacher performance.

The project had one objective related to teacher performance on the MMTsm:

Objective 1.1: By April 29, 2016, all participating teachers will demonstrate at least 10% growth in their mathematical content knowledge, as measured by an approved teacher content assessment (such as the Mathematical Meanings for Teaching secondary mathematics instrument, or MMTsm).

Due to the complex nature of MMTsm scoring and analysis, comprehensive analysis of participant scores was performed by Neil Hatfield, a statistician at Arizona State University who had extensive experience with the instrument. Mr. Hatfield included in his analysis two additional participants who did not complete 90% or more of the professional development (they completed only 88%) but did have matched pre- and post- MMTsm data. His analysis showed that 61% ($n=19$) of the 31 teachers in the experimental group had at least a 10% change in their u^* score and that this change was statistically significant from pre to post. Thus, Project CoMPASS fell short of its target of having all participants demonstrate at least a 10% growth in their mathematical content knowledge. The analysis further showed that, overall, the experimental teachers' improvement in MMTsm scores was not significantly different from that of the control group.

The external evaluator utilized the MMTsm score calculated by Mr. Hatfield to determine the percentage of teachers who completed 90% or more of the professional development who showed significant gain in their content knowledge according to MSP criteria (i.e., using MSPSTCK software). This analysis excluded teachers who completed less than 90% of the professional development. The results of this analysis by the evaluation team showed that 52% ($n=15$) of the 29 teachers in the experimental group who completed 90% or more



of the professional development had a significant gain in their content knowledge from pre to post, which also falls short of the project's target. The complete statistical report for the MMTsm developed by Mr. Hatfield will be submitted as an attachment to the Annual Performance Report.

RTOP

The RTOP consists of five 5-item subscales that measure the degree to which certain student-centered engaged learning practices are present in a classroom. The five subscales are: Lesson Design and Implementation; Propositional Knowledge; Procedural Knowledge; Communicative Interactions; Student/Teacher Relationships; and Lesson Design and Implementation. Each of the 25 items may be scored from 0 to 4, for a maximum total score of 100.

Objective 1.2: By April 29, 2016, all participating teachers will demonstrate at least 10% growth in their propositional pedagogic knowledge, as measured by ratings on the "Content - Procedural Pedagogic Knowledge" subscale on the Reformed Teaching Observation Protocol (RTOP).

Analysis showed that more than half (58%, n=18) of participating teachers achieved a 10% or greater increase in their propositional pedagogic knowledge, as measured by ratings on the Content - Procedural Pedagogic Knowledge subscale. Only 46% (n=13) of the control teachers achieved the same amount of knowledge increase on this subscale.

Objective 3.1: By April 29, 2016, all participating teachers will demonstrate at least 10% growth in their ability to implement formative, in-class assessments and to provide appropriate instruction as measured by ratings on the "Classroom Culture - Communicative Interactions" subscale on the Reformed Teaching Observation Protocol (RTOP).

Analysis also showed that more than half (58%, n=18) of participating teachers achieved a 10% or greater increase in their ability to implement formative, in-class assessments and to provide appropriate instruction as measured by ratings on the Classroom Culture - Communicative Interactions subscale. Only 39% (n=11) of the control teachers achieved the same amount of ability increase on the subscale.

To further inform the evaluation, the evaluation team compared experimental and control teachers' total RTOP scores. The analysis found that 65% of experimental teachers increased their total RTOP scores from pre- to post- compared to only 46% of control teachers. However, this difference was not statistically significant ($\chi^2=1.953$, $df=1$, $p=.196$).



Use of District Benchmark Data

Project CoMPASS has one objective related to participant review and use of district benchmark data.

Objective 3.2: By April 29, 2016, all participating teachers will use district benchmark assessment results to adjust curriculum and instruction.

Limited information is available about the teachers' use their respective district's student benchmark data to adjust curriculum and instruction. FWUSD's middle school and high school math departments, which included all Project CoMPASS participants, met following each round of benchmarks to analyze results and determine next steps. However, teachers made limited adjustments to their curricula and instruction due to a sense of urgency to move forward with completing those curricula. Additionally, students did not perceive a purpose for the benchmark tests given that teachers did not use the outcomes to drive instruction, nor did the tests count toward students' grades. As a result, students underperformed in the assessments. Data are not available on how SUSD and MUSD participants made use of district benchmark data.

Development of Teaching Resources

One of the project's objectives tasked participants with developing professional development resources to be shared online with other educators.

Objective 4.1: By April 29, 2016, all participating teachers will have developed supporting activities, lessons, and formative assessments congruent to the mathematical content addressed throughout the professional development; each participating grade level or course will have developed at least 10 lessons and supporting assessments and posted all resources in a shared online environment.

Participants worked in grade level or course groupings during professional development sessions to create a total of 50 lessons. One the IHE instructors posted the lessons developed in an online Google Group shared website. Exhibit 3 shows the breakdown of the lessons developed.

Exhibit 3. Lesson Developed by Participants

Grade or Subject of Lessons	Number of Lessons
Grade 7	8
Grade 8	11
Algebra 1	10
Geometry	10
Algebra 2/Pre-calculus	11



Application of New Pedagogical Strategies

Qualitative data about the project's impact on teachers' classroom practice was available from a project-developed survey that teachers completed at the last session of professional development. The survey included a question that asked teachers to identify the instructional strategies learned during the professional development that they had already implemented in their classroom. Responses to the open-ended question show that by the last session of the professional development in February 2016, teachers were already applying a range of teaching strategies they had learned during Project CoMPASS. The following are some representative responses.

- Task cards to introduce/expand topics in the classroom. Importance of wait time, random selection, 4 square to help all types of learners.
- Giving students time to struggle with a math problem. One really good problem is better than 10 problems.
- Most of what I have used is the exploration type of lesson strategy. I have also used multiple brainstorming strategies to start lessons depending on the level of student. Many of the formative assessments.
- Project CoMPASS has helped me take my lesson plans to a new level of thinking. Aside from implementing more formative assessments I have created my lesson plans to focus on higher level thinking.
- I learned to allow students to struggle and fail on their way to understanding the content being taught in the classroom. I also learned ways to combine topics together to make the lesson more in-depth.
- Small group/big group discussions. Giving students a chance to brainstorm ideas in a smaller setting and then comparing it to all ideas.
- I have students collaborating more. Use more open ended type of problems than I used to and have my students explain what they were thinking more often than I did before.
- Choices on problems based on difficulty level. More group/partner exploration before starting a lesson.
- More think time. More exploratory time. Better leading questions.



Student Findings

As noted above, the Project CoMPASS co-directors and ADE staff overseeing the MSP program decided for a number of reasons that AzMERIT data would not serve as an appropriate measure of student outcomes (Objective 2.1). Therefore, the only quantitative measure of the project's impact on student achievement are district benchmark data.

District Benchmarks

Project CoMPASS teachers taught a total of 3,221 students in the 2015-2016 school year, 1,920 middle school students and 1,301 high school students. The project used district benchmarks for assessing the impact of the project on student achievement, with each participating district having its own form of benchmark and scoring. FWUSD and SUSD used benchmarks developed in coordination with Assessment Technology Incorporated (ATI) (i.e., ATI's Galileo assessment system). The benchmarks mostly consisted of items from ATI's item bank, with a few modifications to meet each district's needs. MUSD used the second semester final exam grade as its benchmark for high school students but was unable to provide benchmark data for its junior high school students. The project had one objective related to the district benchmarks:

Objective 2.2: By April 29, 2016, students in participating teachers' classrooms will demonstrate a 10% increase in the percentage who score proficient (Meets or Exceeds) on 2016 district/school benchmark assessments compared to 2015 benchmark assessments.

Only aggregate student data was available for each subject they taught in 2015 or 2016 (i.e., the average score for all students taught, per teacher and by subject). Therefore, the evaluation is unable to determine if there was a change in the percentage of students who scored proficient on the benchmarks. Only experimental teachers for whom there were 2015 and 2016 benchmark scores for the same subject were included in the analyses (N=20). Some teachers have student benchmark data for more than one subject. FWUSD and SUSD benchmark data for both junior high school and high school students are the percentage of correct answers on district-wide standardized exams. The MUSD benchmark for high school students is the grade on the Semester 2 final exam (MUSD did not provide middle school data.) Given such differentiation in benchmarks the findings are reported for each district separately. Moreover, data is displayed by subject to further identify differences in the impact of the professional development. These results do not support a finding that the professional development contributed to increases in student benchmark scores (see Exhibits 4-6). However, analyses for this outcome were limited by lack of 2015 and/or 2016 data for some teachers who received the professional development. For example, not all of



the teachers taught the same subject in both years. Furthermore, it is important to note that benchmark scores for the two years are for different groups of students.

Exhibit 4. FUSD Participants' Student Benchmark Change

Subject	Number of Teachers with 2015 and 2016 Student Benchmark Data	Number of Teachers Whose Students Showed a 10% Increase on District Benchmark
Math 7	4	1
Math 8	3	1
Algebra	1	0
Geometry	4	2
Algebra II	5	1

(n=11)

Exhibit 5. MUSD Participants' Student Benchmark Change

Subject	Number of Teachers with 2015 and 2016 Student Benchmark Data	Number of Teachers Whose Students Showed a 10% Increase on District Benchmark
Algebra	3	1
Geometry	3	0
Trigonometry	1	0
Calculus	1	0

(n=4)

Exhibit 6. SUSD Participants' Student Benchmark Change

Subject	Number of Teachers with 2015 and 2016 Student Benchmark Data	Number of Teachers Whose Students Showed a 10% Increase on District Benchmark
Math 7	2	0
Math 8	2	0
Algebra	2	0
Geometry	1	0

(n=5)



Teacher Perceptions of Student Reaction to Use of New Instructional Strategies

Qualitative data from the Participant Survey serves as another source of information on how the project impacted students. The survey asked teachers to describe how their students had reacted to their use of new instructional strategies learned during Project CoMPASS. Responses indicate that many students found the new teaching strategies used by their teachers to be challenging; however, in some classrooms students were becoming more engaged and having more fun with lessons. The following are some representative responses:

- They are willing to try more. They have more confidence in math.
- They have come to be more comfortable with new strategies and sometimes push me to use those strategies when I am not using them. They seem to have more fun at times when I use the exploration especially. Brainstorming sometimes takes more push but once also going they have lots of fun.
- The students were EXTREMELY frustrated with my refusal to give them the answers. My students got into the discovery strategy and even researched to ensure the answer was correct.
- I wouldn't say all students but I am seeing more willingness for the students to come up and explain their math thinking and the class being supportive in everyone learning.
- They dig deeper for answers because they know I will not just give them the answer.
- So-so. I teach at-risk students so the "perseverance" aspect is lacking. They do like the "I don't know...yet" that I've taken because they know they can't get out of answering.
- My students do not enjoy the struggle and failure that has occurred. However, it has allowed them to find the importance of trying to solve problems using what they can bring to the table.
- Most of the students are open to new strategies. However, they struggle with basic math skills and motivation. So it is important to find ways to keep them excited about learning math.
- They realize it is okay to struggle with math and not give up (they still want to give up easily).



- Strategies involving open-ended questions has been extensively used, which has allowed resource students the opportunity to listen to their own reasoning and revise it as needed.
- Most of my students have responded to the new strategies. My class has a multitude of SPED and ELL students that have a hard time just trying to figure out what is being asked.
- Everything I've used so far my students have responded to positively. Some types of assessments have been tricky to implement but eventually something beneficial is able to be drawn from the attempt.



Appendix A

Project Compass Participant Survey

Dear Project Compass Participant,

Thank you for your participation in the project. We would appreciate your feedback on the impact of the project and your satisfaction with the project. This information will help us in reporting to our funder and in designing future professional development opportunities for teachers. If you need more room for any answers, please use the other side of the page.

1. What instructional strategies from the professional development you received through Project Compass have you already implemented in your classroom?
2. What instructional strategies that you learned from the professional development but have not used in the classroom do you plan to use in the future?
3. How have your students responded to your use of new instructional strategies learned during Project Compass?
4. What do you see as the challenges to implementing what you learned in the professional development in your classroom?
5. What kinds of support will enable you to implement what you have learned during Project Compass?

Please also answer the questions on the other side of this page.



6. When you think about your mathematic learning during Project Compass, what was most significant?

7. When you think about your pedagogical learning during Project Compass, what was most significant?

8. What was the most difficult part of the professional development for you?

9. How could the instructors improve future professional development similar to Project Compass?

10. Have you had the opportunity to share some of what you learned during Project Compass with teachers who are not participating? If so, what did you share and in what setting (an in-service, over lunch, etc.)?

11. Would you recommend participation in professional development like Project Compass to other mathematics teachers? If yes, why? If no, why not?

Please also answer the questions on the next page.



For Questions 12-15, please circle the answer that is most true for you.

12. Overall, the pace of the instruction during Project Compass was:

Much too fast	Somewhat fast	Just right	Somewhat slow	Much too slow
------------------	------------------	---------------	------------------	------------------

13. BEFORE participating in Project Compass, how confident were you about teaching the content areas covered during the professional development?

Not at all confident	Somewhat confident	Somewhat unconfident	Very confident
-------------------------	-----------------------	-------------------------	-------------------

14. AFTER participating in Project Compass, how confident are you about teaching the content areas covered during the professional development?

Not at all confident	Somewhat confident	Somewhat unconfident	Very confident
-------------------------	-----------------------	-------------------------	-------------------

15. Please rate the quality of the printed materials you received during Project Compass.

Poor	Fair	Good	Excellent
------	------	------	-----------

Thanks for completing the survey!

