ILLINOIS
CPRD | Center for Prevention Research & Development
SCHOOL OF SOCIAL WORK

Middle-Grades Leadership Development (MLD) Project: A U.S. Department of Education Investing in Innovation (i3) Development Grant Final Evaluation Report

August 2018

Nancy Flowers, Dawn M.H. Carpenter, Shahana Begum

This work was supported by the U.S. Department of Education's Investing in Innovation program, through Grant Number U411C130028 to the National Forum to Accelerate Middle-Grades Reform. The opinions expressed are those of the authors and do not represent the views of the U.S. Department of Education.

| 510 Devonshire Dr, Champaign, IL 61820 | (217) 333-3231 | website: publish.illinois.edu/cprd |

ACKNOWLEDGEMENTS

Middle-Grades Leadership Development (MLD) Project: A U.S. Department of Education Investing in Innovation (i3) Development Grant, Final Evaluation Report

Produced for:
National Forum to Accelerate Middle-Grades Reform
P.O. Box 224
Savoy, IL 61874
middlegradesforum.org

Report Prepared by:
Center for Prevention Research and Development
School of Social Work
University of Illinois
510 Devonshire Drive
Champaign, IL 61820
publish.illinois.edu/cprd

Nancy Flowers, Director of Research Programs
Dawn M. H. Carpenter, Senior Research Program Coordinator
Shahana Begum, Research Data Coordinator

This work was supported by the Raikes Foundation; Edmodo; the Association for Middle Level Education (AMLE); the National Association of Secondary School Principals (NASSP); the Kentucky Middle School Association (KMSA); and the Institute for Excellence in Education (IEE) through matching funds or in-kind donations.

TABLE OF CONTENTS

Table of Contents

Note: Throughout this report, references to "project schools," "MLD Project schools," or "intervention schools" mean the schools that received all project services and activities. In contrast, schools that participated in data collection, but did not receive any project services or activities, are referenced only as "comparison schools."

EXECUTIVE SUMMARY

Executive summary

MLD PROJECT OVERVIEW

The Middle-Grades Leadership Development (MLD) Project was designed to develop principal leaders and leadership teams who create high-performing middle-grades schools. Developed by the National Forum to Accelerate Middle-Grades Reform, the four-year project was funded in 2013 by a U.S. Department of Education Investing in Innovation (i3) development grant. The project was implemented in Kentucky by the Kentucky Middle School Association (KMSA) and in Michigan by the Institute for Excellence in Education (IEE). The Center for Prevention Research and Development (CPRD) at the School of Social Work, University of Illinois served as the evaluator.

The MLD Project used an extensive set of school improvement supports, including: creating a vision using the Schools to Watch (STW) criteria; engaging in an assessment and planning process for improvement; STW leadership coach; principal mentor; STW mentor schools; leadership team; networking opportunities; and focused professional development. Project planning began in January 2014 with the national and state partners, and the 12 MLD Project middle-grades schools (6 in Kentucky and 6 in Michigan) began receiving services in July/August 2014 through May 2017.

> "The grant allowed for the implementation of distributive leadership. Through this, we were able to develop a common vision throughout our faculty about what good instruction looks like, sounds like, and how we want our classrooms to function. We changed our professional development model to one where we learn from each other. This allowed a team of leaders to be developed around the school. We truly have become a faculty of leaders."
>
> - MLD Principal

EVALUATION DESIGN

The evaluation of the MLD Project used a quasi-experimental design (QED) with matched comparison schools. Twelve middle-grades schools in Kentucky and Michigan (6 schools per state) received the project treatment. Each of the twelve treatment schools were matched with two to four comparison schools from the same state and with a similar profile of achievement test scores, No Child Left Behind (NCLB) Adequate Yearly Progress (AYP) status and history, and school and student demographics. A cohort of 6th graders from the 50 schools (12 treatment and 38 comparison) were followed for three years, and math and ELA/reading state achievement test scores (Kentucky Performance Rating for Educational Progress and Michigan Student Test of Educational progress) were the primary outcomes, along with principal effectiveness scores via Vanderbilt's *Assessment of Leadership in Education* (VAL-ED). Intermediate outcomes were tracked

EXECUTIVE SUMMARY

with mixed methodologies consisting of data collected via the National Forum's STW *Rating Rubric*, CPRD's *School Improvement Self-Study Teacher Survey*, focus groups with coaches and principals, a principal survey, and CPRD's *Coach's Log*. The project's impact – defined as the difference between the project and comparison schools on achievement test scores – was estimated using a hierarchical linear model (HLM) approach. The overall intervention effect between the project and comparison schools on achievement test scores was examined.

MLD PROJECT PRINCIPALS

The 12 MLD Project principals were experienced educators with an average of 17.6 years of work in the field of education and an average of 13.2 years working in the middle grades. Principals reported an average of 8.5 years serving as a principal with 5.1 of those years as a principal in their MLD Project school. All MLD Project school principals had a school administrator/principal certification. A total of 8 principals had a Master's degree in Education and the remaining 4 had an Educational Specialist degree. Although very few received pre-service specialized training in the middle-grades education, all participated in professional development on middle-grades instruction and organizational needs during their years as a teacher or administrator. There was collective agreement that the MLD Project was helping to improve leadership at their school, both for themselves as well as for their faculty

IMPLEMENTATION FINDINGS

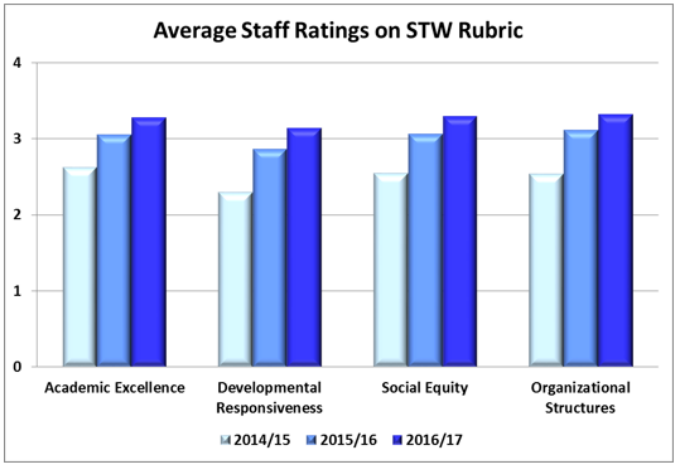
The eight key programmatic components (create a powerful vision using STW; engage in assessment and planning to identify needs, develop goals, and implement an action plan; STW leadership coach; principal mentor; STW mentor school; leadership team; networking opportunities; and focused professional development) of the MLD Project were implemented with fidelity at the majority of project schools. There was some variability in implementation by school, with four schools having slightly higher and more consistent overall implementation. In focus group interviews with coaches, they commented that several schools had a higher level of readiness for making improvements at the start of the project and their districts were highly involved and supportive of their goals.

INTERMEDIATE OUTCOME FINDINGS

Evidence indicates that MLD Project schools significantly improved their implementation of the STW Rubric criteria for high performing middle-grades schools, collaboration practices, teacher efficacy, and middle-grades classroom instructional practices.

EXECUTIVE SUMMARY

- **STW Rubric** – Higher implementation of strategies to support academic excellence, developmental responsiveness, social equity, and organizational structures.
- **Collaboration practices** – Increased team practices, quality of collaborative interactions, and team decision making.
- **Teacher efficacy and collective responsibility** – Improvements in teachers' individual and collective commitment to the success of their students.



4=High quality; 3=Good quality; 2=Fair quality; 1=Poor quality p≤.001

- **Middle-grades instructional practices** – Increased recommended practices such as small group instruction, integration and interdisciplinary practices, and critical thinking practices.

FINAL OUTCOME FINDINGS

MLD Project principals improved their leadership skills and behaviors throughout the project, with principals demonstrating the greatest growth in the areas of culture of learning and professional behavior, quality instruction, performance accountability, and high standards for student learning. Overall principal effectiveness, as measured by all respondent groups (i.e., principal, teachers, supervisor), improved from two principals rated as proficient or distinguished in Year 1 of the project to nine principals by the end of Year 3. Based on the percentile score of the principal's mean score in comparison to a national sample of principal means, all twelve principals grew as leaders when compared to the national sample.

> "The overall impact of the grant has been very positive. It has given us purpose, direction, and guidance to improve what we do for the overall growth of the whole student."
>
> - MLD Principal

A comparison of the MLD project schools to the state at the 8th grade level in 2017 (after the three-year project implementation) showed that seven schools had larger reading growth than the state and three schools had larger math growth than the state. Although these are positive trends in achievement, further analyses were conducted on the cohort of students tracked for

three years at MLD Project schools as compared to a cohort at comparison schools using Hierarchical Liner Models (HLMs). The series of 2-level models, where students were nested within schools, were executed to analyze the intervention effect on student achievement. The final models suggest no significant intervention effect on either ELA/reading scores or math scores. In other words, the students that received the MLD Project intervention performed the same on both the ELA/reading test and the math test as the comparison students after three years. Neither the p-values nor effect sizes suggest a significant intervention effect on achievement. Even though the impact analyses did not find an overall intervention effect on achievement, it should not be interpreted to mean that the MLD Project intervention was not effective, but that project schools may need additional time to fully implement and refine the model with middle-grades students to advance achievement scores. We can further hypothesize that the positive changes on school organization and climate factors such as the STW criteria, collaboration, teacher efficacy, middle-grades instructional practices, and leadership are a positive precursor to improvements in achievement since achievement does not improve without corresponding improvements in the teaching and learning environment.

LESSONS LEARNED

There are a myriad of lessons learned from the MLD Project about how middle-grades schools embark on leadership improvements and the implementation of a collaborative leadership model, including:

- The STW criteria provided a guiding vision to MLD Project schools for what a high performing middle-grades school should be. It served as a framework, common language, and pathway to improvement.
- The adoption and use of a continuous school improvement model where data was used at every stage of a cycle to inform planning, set targeted goals, reflect on and evaluate progress, and refine implementation, was a key aspect remain focused on improvement.
- The implementation of a collaborative leadership structure empowered teachers to take on shared leadership roles and build capacity.
- The development of strong leadership teams that were engaged in collaboration and focused on improvement was a key to facilitating change. Leadership teams examined data to drive actions, had a continuous improvement approach, communicated with the whole faculty, built capacity for sustainability, and evolved as they learned from challenges.

EXECUTIVE SUMMARY

- Middle-grades principals need support and resources to assist them with transforming their leadership practices, beliefs, and behaviors.
- Visits to STW schools made a powerful impact because they allowed teachers to observe best practices, gain knowledge about the successes at another school, and share a common experience.

> "I think for me the biggest piece was just to increase my knowledge about what works in the middle grades and how we can best implement those practices in my school setting."
>
> - MLD Principal

- Participation in the Forum's STW network at the state and national levels provided schools with knowledge, resources, and supportive professional development opportunities.
- It was important to build a positive, supportive, reflective, and student-centered school culture. The contextual changes in culture were viewed as pre-requisites for leadership change, collaboration opportunities, and instructional improvements.
- Cultivating school district involvement and support of the school's improvement plans from the beginning and maintaining it throughout the project was an important part of successful improvements.

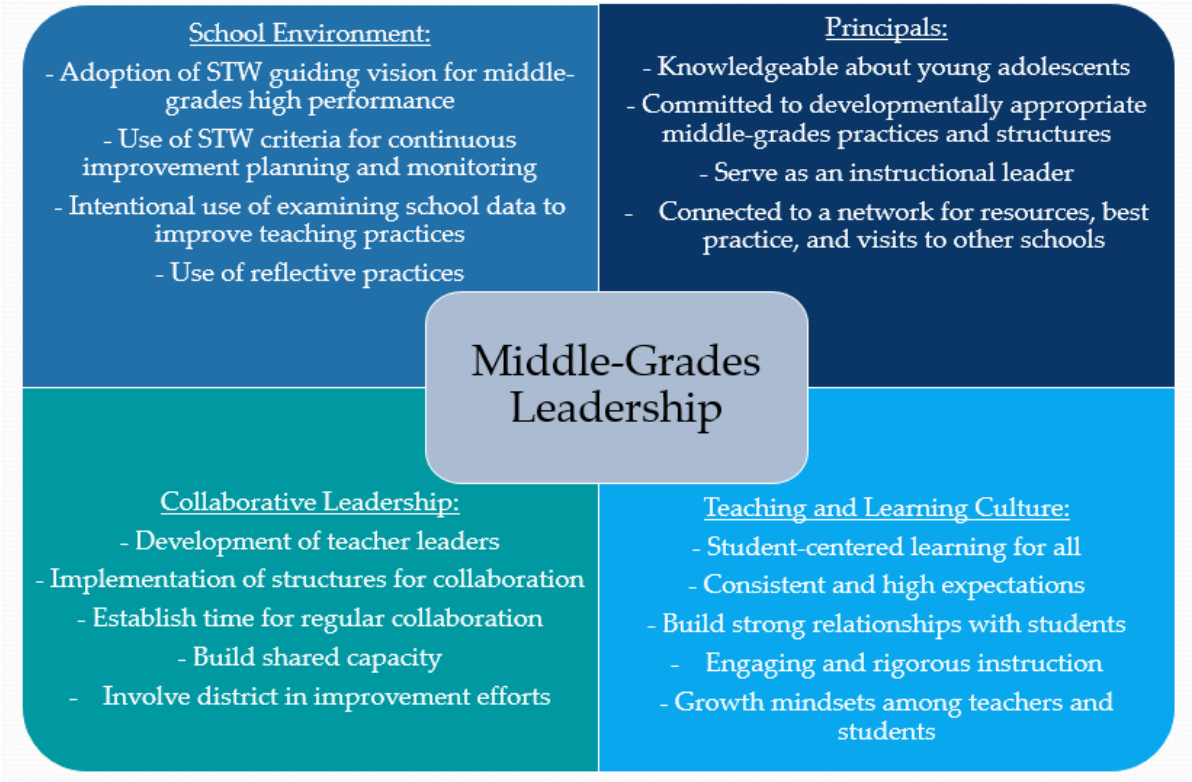
> "We're moving from an adult-centered to a student-centered facility and programs. The piece that has really bought us growth from a teaching standpoint is active engagement strategies for youngsters in the classroom and that is really elevating the teaching and learning. That is critical. They absolutely engage kids. We've got a group of kids that have to be active, so might as well channel it in a useful way."
>
> - MLD Leadership Coach

- Participation in mock STW visits was the activity that taught schools the most about their improvement progress and goals. Mock STW visits involved the MLD Project schools hosting a national team of STW-trained visitors whose purpose was to spend a day touring the school, meeting with administrators and teachers, observing classrooms, and interviewing students and parents, then providing feedback to the school leadership on future direction.

CONCLUSION

The results of the evaluation of the MLD Project are positive in that they highlight numerous gains and improvements at MLD Project schools. The results also provide unique insight into a middle-grades school improvement project focused on principal leaders and collaborative leadership. The figure below depicts the key supports, activities, and practices implemented at MLD Project schools that were the most impactful on building middle-grades leadership effectiveness. They provide a road map for other middle-grades schools and principals that are struggling to improve. With key supports, activities, and practices at the school level, by

the principal, through a collaborative leadership structure and in a positive teaching and learning culture, middle-grades leadership is more effective.



MLD Project

The Middle-Grades Leadership Development (MLD) Project was designed to develop principal leaders and leadership teams who create high-performing middle-grades schools. Developed by the National Forum to Accelerate Middle-Grades Reform, the four-year project was funded in 2013 by a U.S. Department of Education Investing in Innovation (i3) development grant. The project was implemented in Kentucky by the Kentucky Middle School Association (KMSA) and in Michigan by the Institute for Excellence in Education (IEE). The Center for Prevention Research and Development (CPRD) at the School of Social Work, University of Illinois served as the evaluation partner for the MLD Project.

The MLD Project sought to improve principal effectiveness in order to improve the academic achievement of high-needs, middle-grades students at 12 schools, through the development of principal and leadership team skills and behaviors to create high-performing, middle-grades schools and through improved climate and culture for learning.

> "It's really easy for everybody to focus on academic achievement. Everything is driven by test scores. I think teachers have really appreciated having the four areas spelled out in the STW Rubric criteria, knowing that in middle school it is so important to be in tune with student development."
>
> - MLD Principal

The intervention intended to accomplish this goal using the framework of the National Forum to Accelerate Middle-Grades Forum's Schools to Watch (STW) vision and criteria with the following program objectives:

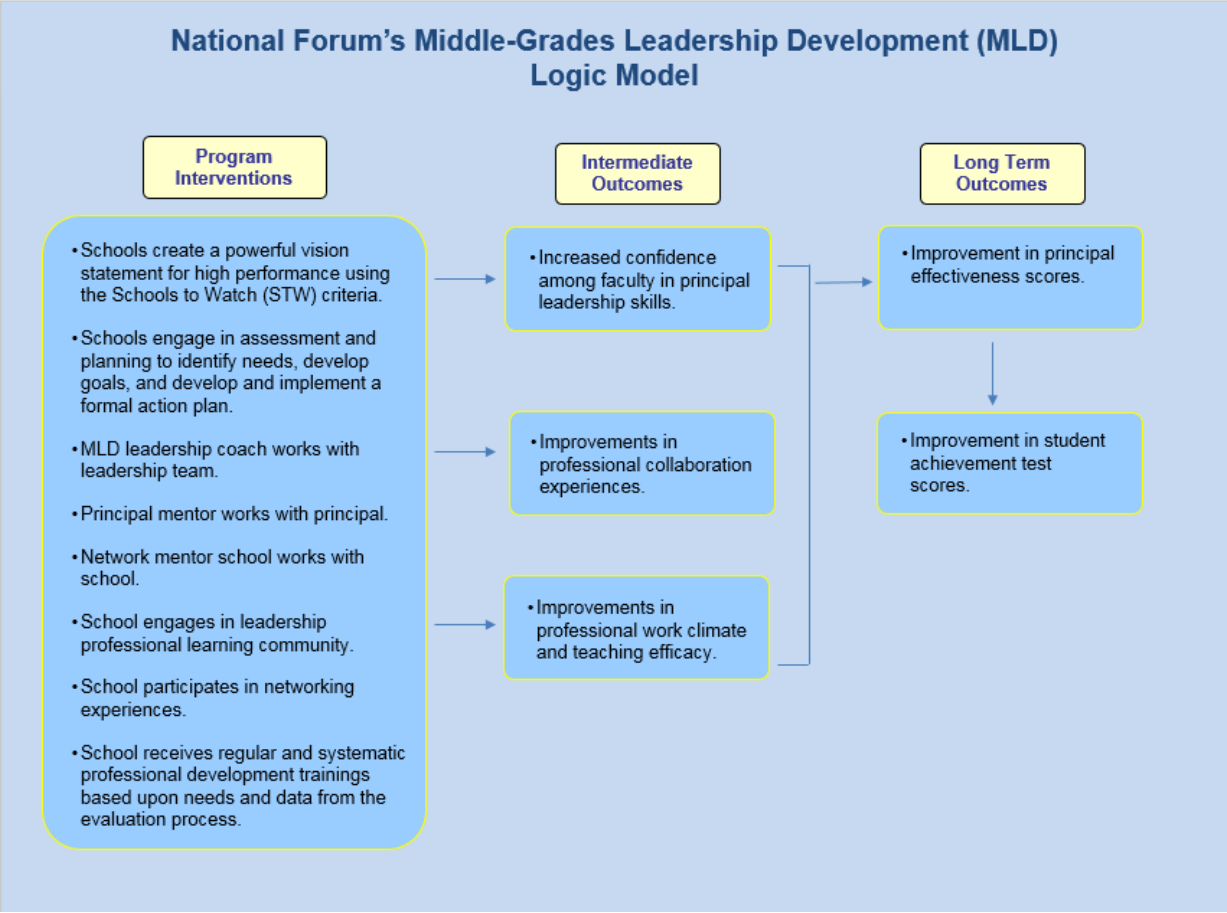
1) Improve principal and leadership team skills, behaviors, and practices for continuous improvement;
2) Improve school implementation of the four key criteria of the STW vision (academic excellence, developmental responsiveness, social equity, and organizational structures and processes) by implementing programs, practices, and organizational structures consistent with the vision;
3) Improve school climate and culture for learning by engaging in the STW network to exchange ideas, solve problems, and discuss practice; and
4) Improve student achievement outcomes.

Project planning and development began in January 2014 with the national and state partners. The 12 MLD Project middle-grades schools (6 schools in Kentucky and 6 schools in Michigan) began receiving services in July/August 2014.

INTERVENTION ACTIVITIES/KEY PROGRAM PIECES

The MLD Project used an extensive set of school improvement supports, including a Schools to Watch (STW) leadership coach, principal mentor, STW mentor schools, focused professional development, and networking opportunities. Through these supports and activities, schools created a powerful vision for high performance using the STW criteria; engaged in an assessment and planning process to develop and implement a formal action plan; and used intervention supports and services to develop leadership capacity and sustainability in professional learning communities. The project was guided by a logic model (Figure 1).

Figure 1. MLD Project Logic Model



EXPECTED OUTCOMES

It was hypothesized that schools engaged in the MLD Project would experience improvements in the following intermediate outcomes: increased confidence among faculty in the

principal's leadership skills; improvements in professional collaboration experiences; and improvements in professional work climate and teaching efficacy. Ultimately, success would be measured on the long-term outcomes of the project: improvements in principal effectiveness scores and improvements in student achievement test scores.

Principal effectiveness was measured with the VAL-ED Leadership Assessment. Student achievement was measured with state standardized achievement tests in math and ELA/reading (K-PREP test in Kentucky, M-STEP test in Michigan).

PROJECT SCHOOLS

The twelve schools that participated in the MLD Project were located in primarily rural or small town areas of Kentucky and Michigan, each serving an average of 424 students (Table 1). The average percentage of students receiving free or reduced-priced lunch across the 12 schools was 63%. Approximately 18% of students were minorities and only 2% were English language learners.

Table 1. MLD Project School Demographics

Schools	Districts	City, State	Grade Levels	Enroll-ment	Percent Free/ Reduced Lunch	Percent Minority	Percent English Language Learners
Estill County Middle School	Estill County Schools	Irvine, KY (Rural)	6-8	542	72%	1%	0%
Garrard Middle School	Garrard County Schools	Lancaster, KY (Rural)	6-8	545	63%	10%	10%
Georgetown Middle School	Scott County Schools	Georgetown, KY (Suburban)	6-8	485	57%	21%	1%
Mayfield Middle School	Mayfield Independent Schools	Mayfield, KY (Small Town)	6-8	338	80%	18%	2%
Rowan County Middle School	Rowan County Schools	Morehead, KY (Small Town)	6-8	737	62%	7%	<1%
Union County Middle School	Union County Schools	Morganfield, KY (Rural)	6-8	504	59%	15%	0%
Kentucky Totals				3,151	64%	12%	2%
Grant Middle School	Grant Public School District	Grant, MI (Rural)	5-8	430	55%	25%	8%
Marshall Greene Middle School	Birch Run Area Schools	Birch Run, MI (Rural)	5-8	408	50%	9%	2%

Mt. Morris Junior High School	Mt. Morris Consolidated Schools	Mt. Morris, MI (Suburban/ Urban)	6-8	430	75%	23%	<1%
Oscoda Schools	Oscoda Area Schools	Oscoda, MI (Rural)	K-12	274	52%	2%	<1%
Reese Middle School	Reese Public Schools	Reese, MI (Rural)	5-8	176	47%	10%	2%
Richfield Public School Academy	Richfield Public School Academy	Flint, MI (Urban)	3-8	223	90%	75%	7%
Michigan Totals				1,941	63%	24%	3%
Project Totals				5,092	63%	18%	2%

Notes: Data from 2014/15 (Year 1). Enrollment includes only grades served (6-8).

Evaluation Design

The Center for Prevention Research and Development (CPRD) at the School of Social Work, University of Illinois served as the evaluation partner for the MLD Project. The evaluation used a quasi-experimental design (QED) with matched comparison schools. Twelve middle-grades schools in Kentucky and Michigan (6 schools per state) received the project treatment. Each of the twelve treatment schools were matched with two to four comparison schools from the same state and with a similar profile of achievement test scores, No Child Left Behind (NCLB) Adequate Yearly Progress (AYP) status and history, school enrollment size, and school level demographics. A cohort of 6th graders from the 50 schools (12 treatment and 38 comparison) were followed for three years, and math and ELA/reading state achievement test scores were the primary outcomes. Intermediate outcomes were also tracked at intervention schools with mixed methodologies consisting of data collected via Vanderbilt's *Assessment of Leadership in Education* (VAL-ED), the National Forum's STW *Rating Rubric*, CPRD's *School Improvement Self-Study Teacher Survey*, focus groups with coaches and principals, a principal survey, and CPRD's Coach's Log. The data analysis plan for assessing the impact of the project used a hierarchical linear model (HLM) approach. The overall intervention effect between the project and comparison schools on achievement test scores was examined.

SAMPLE

Fifty public middle-grades schools serving grades 6th to 8th in Kentucky and Michigan (6 treatment and 19 comparison schools per state) participated in the evaluation. All schools served high-need student populations with suppressed achievement test scores. A cohort of 6th graders were followed at treatment schools for three years, comparing their outcomes to a cohort at comparison schools. A total of 1,304 treatment and 4,433 comparison students participated in the cohort. Comparison schools were "business as usual" middle-grades schools that did not receive the MLD Project intervention, i.e., levels of support, training, or mentoring for their principals and school leadership.

Within Kentucky and Michigan, a pool of potentially participating middle-grades schools was developed. The treatment sample was selected using detailed criteria that each school must meet in order to be considered for inclusion in the project, including: 1) Public school; 2) Middle-grades school with a grade configuration of at least 6th through 8th; 3) Serve high-need student populations (e.g., students at risk of educational failure, such as students who are living in poverty, who are English language learners, who are performing below grade level); and 4) Suppressed achievement scores. This criteria was used to identify a pool of eligible schools and recruit their participation. It is considered a convenience sample with well-defined criteria.

EVALUATION DESIGN

The comparison schools were selected from the same pool of eligible schools used to select the treatment schools. The matching was conducted using propensity scores at the school level within each state from the pool of eligible schools. Schools were matched on the following variables: state, enrollment, student demographics, achievement test scores, and NCLB AYP status and history. When possible, matches were explored within the same district as treatment schools. Schools were matched first on school level achievement test scores and second on NCLB AYP status and history. Once achievement test scores and NCLY AYP status and history were matched, schools were matched on the remaining variables in the list above. For some schools, matches were difficult on the remaining list of multiple variables, thus we allowed a 5% to 10% margin of difference as the target. Baseline equivalence tests between the treatment and comparison schools were conducted based on their 5th grade achievement scores (i.e., the year before the intervention began) and other available demographic variables on the final analytic sample to verify the equivalence between treatment and comparison schools (see pages 36-37 for more information). Treatment schools appeared similar compared to the comparison schools based on their 5th grade achievement scores as well as the other student level demographic variables. Yet to reduce the student level variances, the baseline scores and other covariates were used to adjust the evaluation models.

All students enrolled in treatment schools and comparison schools as 6th graders in 2014/15 were part of the cohort of students that was followed for three years. Students who entered the schools later in 2014/15 or in between project years were excluded. Students with no pretest score (5th grade) were also excluded. A total of 1,304 6th graders from twelve treatment schools and 4,433 6th graders from the thirty-eight comparison schools over a three year period were tracked.

RESEARCH QUESTIONS

The purpose of the impact study was to determine if the MLD Project intervention significantly improved student math and ELA/reading achievement after three years of exposure, relative to comparison students. There are two confirmatory research questions and one exploratory research question for the impact study.

Impact Study Confirmatory Research Questions:

1) Did the MLD Project intervention have an effect on the math achievement of 8th grade students, after three years of intervention, as compared to business-as-usual condition?

2) Did the MLD Project intervention have an effect on the ELA/reading achievement of 8th grade students, after three years of intervention, as compared to business-as-usual condition?

EVALUATION DESIGN

Impact Study Exploratory Research Question:

3) Did the MLD Project intervention improve principal leadership practices, organizational capacity, and school influences?

The purpose of the implementation study was to document the fidelity of implementation of the MLD Project at each treatment school, assessing whether the intervention was implemented as intended. There is one exploratory research question for the implementation study.

Implementation Study Exploratory Research Question:

1) Was the MLD Project intervention implemented with fidelity at each treatment school?

DATA SOURCES

Data sources for the evaluation included mixed methodologies. Math and ELA/reading achievement test scores were collected for a cohort of 6th graders from the 50 schools (12 treatment and 38 comparison) who were followed for three years (Table 2). Intermediate outcomes were tracked at treatment schools with data collected via Vanderbilt's *Assessment of Leadership in Education* (VAL-ED), the National Forum's STW *Rating Rubric*, CPRD's *School Improvement Self-Study Teacher Survey*, focus groups with coaches and principals, principal survey, and CPRD's *Coach's Log.*

Math and ELA/reading achievement test scores were collected from each state. For project and comparison schools in Kentucky, the Kentucky Performance Rating for Educational Progress (K-PREP) achievement test was used. For project and comparison schools in Michigan, the Michigan Student Test of Educational Progress (M-STEP) test was used. Both represent the state approved standardized achievement test. The achievement data was obtained from the state departments of education via a data sharing agreement during the spring of each project year, beginning with the baseline 5th grade test scores in 2014.

VAL-ED (Vanderbilt Assessment of Leadership in Education) is a research-based evaluation tool created in 2006 by Vanderbilt University (with support from the Wallace Foundation and the U.S.D.E. Institute for Education Sciences) that measures the effectiveness of school principals using 360 degrees of feedback from teachers, supervisors, and principals. The VAL-ED assesses principals in the following six core components of leadership and six key processes of leadership:

- **Core Components of Leadership:** high standards for student learning, rigorous curriculum, quality instruction, culture of learning and professional behavior, connections to external communities, and performance accountability; and
- **Key Processes of Leadership:** planning, implementing, supporting, advocating, communicating, and monitoring.

The valid and reliable results of the VAL-ED Assessment allow the principal to compare their ratings on each key component or process against the ratings given by teachers and the supervisor. In this way, the principal receives informative feedback about their leadership behaviors in order to develop a professional growth plan. Additionally, the report contained norm-referenced and criterion-referenced scores for assessing leadership. VAL-ED was administered at project schools annually.

STW Rating Rubric is a tool developed by the National Forum in 1998 that measures the implementation of the criteria associated with the four components of the STW Program. The rubric is used by middle grades schools to study and rate their practices as part of a continuing improvement process as well as part of a mandatory self-rating for schools interested in applying for a STW designation. The four STW program components measured by the rubric include:

1. Academic Excellence – Schools challenge students to use their minds well;
2. Developmental Responsiveness – Schools are sensitive to the unique developmental challenges of early adolescence;
3. Social Equity – Schools are democratic and fair, providing every student with high-quality teachers, resources, learning opportunities and supports; and
4. Organizational Structures and Processes – Schools establish norms, structures, and organizational arrangements to support and sustain their trajectory toward excellence.

For each of the four STW program components, teachers at treatment schools rated their school's level of implementation by responding to 37 general criteria and 100 concrete examples of excellence. Teachers used a metric ranging from one to four where: 4 = High quality, complete, mature, and coherent implementation – NEARLY PERFECT, LITTLE ROOM FOR IMPROVEMENT; 3 = Good quality, maturing but not fully implemented by all – GOOD QUALITY BUT STILL ROOM FOR REFINEMENT AND IMPROVEMENT; 2 = Fair quality, mixed implementation, immature practice, sporadic by some – SIGNIFICANT IMPROVEMENT NEEDED; and 1 = Poor quality, low level of implementation, new program – CONSIDERABLE PLANNING, CONSENSUS BUILDING AND IMPROVEMENT NEEDED. The STW Rating Rubric was administrated at project schools annually.

EVALUATION DESIGN

School Improvement Self-Study Teacher Survey is part of a data collection system of surveys for middle-grades schools, developed by CPRD in 1990. The surveys are grounded in research and have been used with more than a thousand schools, with results widely disseminated (Flowers & Mertens, 2003; Flowers, Mertens, & Mulhall, 1999, 2000a, 2000b, 2002, 2003, 2007; Mertens & Flowers, 2003, 2006; Mertens, Flowers, Hesson-McInnis, & Bishop, 2006, 2007). The teacher survey is comprised of numerous constructs related to the teaching and learning process including: interdisciplinary teaming practices, quality of team interactions, team decision making, work climate, collective teacher efficacy, teacher decision making, administrative leadership, and classroom instructional practices. The items that make up each construct on the teacher survey were combined and scale scores were calculated based on Cronbach's alpha (.76 to .96). The Self-Study Teacher Survey was administrated at project schools annually.

Focus Groups were conducted with a sample of project principals and leadership coaches annually in order to ensure that the voice of each principal and coach was represented over the course of the MLD Project. The purpose of the focus groups was to provide formative results regarding the experiences of implementation and to assess the multi-layered system of support (e.g., coaching, mentoring activities, networking, etc.). Focus group topics changed each year to address the different stages of implementation and to answer questions not addressed by other data sources. Key areas discussed included leadership development (i.e., principal, leadership team, teacher leaders), barriers to implementation and lessons learned, assessing which program interventions had the greatest impact, building capacity for sustainability, and the pathway to STW School designation.

A Principal Survey was conducted with all treatment school principals during the final year of the grant. The purpose of the survey was to collect information about each principal's experiences in education and the middle grades, the helpfulness of the grant services they received, and the impact the grant had on their school. All twelve principals completed the survey. These data were used in analyses in order to disaggregate data by principal experience and certification in order to understand variances in outcomes across the treatment schools who participated in the MLD Project.

Coach's Log is an electronic data collection system for leadership coaches to complete after each visit or activity with a project school. The purpose of the log was to document the number of visits made to schools, the purpose of the visit (e.g., training, professional development, reviewing data, etc.), the outcome of the visit (e.g., goals set, action plan distributed, etc.), as well as to log the improvement progress of the school. The coach's log was completed by leadership coaches on an ongoing basis throughout the MLD Project.

EVALUATION DESIGN

Table 2. MLD Project Data Components and Sample Sizes

Data Component	School Year 1 (2014/15)	School Year 2 (2015/16)	School Year 3 (2016/17)
Achievement Test	1,304 project students & 4,433 comparison students tracked for 3 years		
VAL-ED	300 assessments	292 assessments	279 assessments
STW Rubric	355 staff	347 staff	322 staff
Teacher Survey	323 teachers	270 teachers	288 teachers
Coach's Log	191 Total Contacts	268 Total Contacts	256 Total Contacts
Focus Groups	7 Principals; 6 Coaches	6 Principals; 5 Coaches	6 Principals; 6 Coaches

Feedback reports were disseminated annually for all data sources to the various project stakeholders and to treatment schools. Aggregated achievement data was reported for groups of students (i.e., grade, school, and cohort-level data) to state leaders and project leadership team to monitor the impact of the intervention. Reports from all four of the process measurement tools (i.e., STW Rubric, Self-Study Teacher Survey, Coach's Log, and Focus Groups) were disseminated annually to project schools, leadership coaches, state leaders, and the project leadership team for use in monitoring improvement progress, setting goals, and refining the intervention. School, state, and project-level feedback reports were provided for the STW Rubric and Self-Study Teacher Survey. State-level reports were provided for the Coach's Log. Focus group reports were aggregated across all participants. Longitudinal data was provided for every year of participation in the STW Rubric, Self-Study Survey, and Coach's Log.

DATA ANALSIS APPROACHES

The 2014/15 school year was the first year that the MLD Project was implemented in treatment schools, and the treatment extended through spring 2017. Treatment and comparison students in sixth grade in 2014/15 were followed (Table 3) through spring 2017 (the end of their 8th grade year). Math and ELA/reading achievement state standardized test scores were collected in spring 2015, 2016, and 2017. The 2017 scores served as the outcomes for the confirmatory research questions.

Table 3. MLD Project Student Cohort Tracked

	School Year 1 (2014/15)	School Year 2 (2015/16)	School Year 3 (2016/17)
Cohort Tracking	6th Grade	7th Grade	8th Grade

EVALUATION DESIGN

For the impact study, hierarchical linear models (HLMs), clustering students within schools, were used to analyze the intervention effect on student achievement scores. The dependent variables were related to two outcome domains: 8th grade achievement test scores in mathematics and 8th grade achievement test scores in ELA/reading, after following the cohort of 6th grade who continued to belong to their respective treatment schools and hence receive three years of the intervention. For schools in Kentucky, the Kentucky Performance Rating for Educational Progress (K-PREP) achievement test was used. For schools in Michigan, the Michigan Student Test of Educational Progress (M-STEP) test was used. Both are standardized tests and hence there are no concerns regarding reliability, over-alignment with the intervention, or inconsistent data collection. However, to allow comparison of scores across different state tests and over years, each state's achievement scores were converted into z-scores (standardized to the mean and SD of the state test for each year for that subject) using the test technical manual. The z-score formula used was: $Z = \frac{x - \mu}{SD}$ where x is the student's scale score, μ is the corresponding state mean, and SD is the corresponding state standard deviation matched for the corresponding year.

To evaluate the impact of the intervention, an independent variable was generated that identifies the treatment and comparison schools coded as 1 and 0, respectively. The following covariates were included as part of adjusting for baseline differences with the goal of improving the accuracy of the coefficients and to be aligned with the covariates used in propensity scores matching while selecting the comparison schools.

Student-level covariates – 5th grade pre-test achievement scores, race/ethnicity (white vs. not-white, African American vs. not African American, Hispanic vs. non-Hispanic), gender, LEP status, free/reduced lunch status, and special education status; and

School-level covariates – free/reduced lunch, enrollment, NCLB AYP status, and school demographics.

To estimate the mean baseline difference, the same modeling approach was used for estimating differences on outcomes, but without any covariates. For the implementation study, each key project component was scored using multiple indicators. An implementation score was determined for each indicator. All indicators except one were measured at the school level. The one indicator not measured at the school level was measured at the teacher level instead. Once the indicator level scores were calculated for each project component, the individual indicator scores were scored at the key project component level to arrive at a component level implementation score. Finally, for each component, a threshold was established for assessing fidelity of implementation across the sample (e.g., 85% percent of schools must have high implementation) in order to determine if fidelity of each component was met.

MLD Project Principals

In the final year of the grant, the evaluation team requested that each MLD Project principal complete a short survey. The survey asked about their experiences in education and the middle grades, the helpfulness of the grant services they received, and the impact the grant had on their school.

PRINCIPAL PROFESSIONAL EXPERIENCES

The MLD Project principals were experienced educators with an average of 17.6 years of work in the field of education and an average of 13.2 years working in the middle grades (Table 4). Principals reported an average of 8.5 years serving as a principal with 5.1 of those years as a principal in their MLD Project school.

Table 4. MLD Project Principal Professional Experiences

Professional Experiences	Average Number of Years
Length of time worked in the field of education	17.6 years
Length of time worked in the middle grades	13.2 years
Length of time as a principal	8.5 years
Length of time as a principal at this MLD Project school	5.1 years

PRINCIPAL CERTIFICATION AND PROFESSIONAL DEVELOPMENT

All MLD Project school principals have a school administrator/principal certification. A total of 8 principals have a Master's degree in Education and the remaining 4 have an Educational Specialist degree. All indicated that their major field of study for their highest degree was in education (i.e., educational leadership, instructional leadership, school administration). When asked what teaching certification they had received, they were split between elementary, secondary, and middle grades (Table 5). Eleven of the twelve principals had prior experience as an assistant principal; eight of the twelve had been a middle school teacher as well (average was 10 years as a teacher).

MLD PROJECT PRINCIPALS

Table 5. MLD Project Principal Teaching Certification

Teaching Certification	Number of Principals*
Elementary certification	7
Secondary subject-matter certification	6
Middle grades certification	1
Middle grades endorsement	5
Other (e.g., administrative)	6

* Number of principals does not total 12 because principals may have selected more than one response.

When asked about whether they had received any specialized training in the instruction and organizational needs of a middle school, either during pre-service education or during professional development after joining the education workforce, the largest number of principals received such training while working as a teacher or administrator (Table 6).

Table 6. MLD Project Principal Middle School Training

Specialized Training on Middle Schools	Average Amount*
Pre-service coursework	1.7
Certification coursework	1.9
Professional development	2.6
Master's degree	2.2
Doctoral degree	1.0

* 1=None; 2=1-2 courses; 3=3-5 courses; 4=6 or more courses.

PRINCIPAL RATINGS OF GRANT ACTIVITIES AND THEIR IMPACT

When principals were asked about the impact of the grant on their school, there was collective agreement that it was helping to improve leadership at their school, both for themselves as well as for their faculty (Table 7). Principals also overwhelmingly agreed that the

"I have grown as an instructional leader. My coach provided opportunities for me to see strong leadership. Our teachers have been able to go into other schools and bring back activities and strategies that work with our students. Our school environment has improved." - MLD Principal

improvements they made and the new practices they had implemented would continue after the grant ends.

Table 7. MLD Project Principal Reports of Impact of MLD Project

Impact of the MLD Project	Average Agreement*
There is collective support among the faculty for the implementation of the MLD Project.	4.2
My school's involvement with the MLD Project is helping to improve leadership at my school.	4.5
There is strong commitment among the faculty to support the leadership of my school.	4.3
Working with the leadership coach helped me improve my leadership.	4.4
Working with the leadership coach helped the faculty to improve their leadership.	4.2
Working with the leadership coach helped teachers improve middle-grades practices.	4.2
The MLD Project has improved student learning/achievement at my school.	4.1
My school's participation on a STW designation team was helpful for our continuous school improvement process.	4.3
The improvements our school has made during the MLD Project will continue after the grant ends.	4.5
I will continue to implement the i3 MLD practices and techniques I learned after the grant ends.	4.5

* 1=Strongly disagree; 2=Disagree; 3=Neither agree nor disagree; 4=Agree; 5=Strongly agree.

When principals were asked about their level of satisfaction with the MLD Project services and activities, they indicated they were most satisfied with the leadership coach working with them, the leadership coach working with their leadership team, and hosting a mock STW visit at their school (Table 8).

"The grant allowed for the implementation of distributive leadership. Through this, we were able to develop a common vision throughout our faculty about what good instruction looks like, sounds like, and how we want our classrooms to function. We changed our professional development model to one where we learn from each other. This allowed a team of leaders to be developed around the school. We truly have become a faculty of leaders."

- MLD Principal

MLD PROJECT PRINCIPALS

Table 8. MLD Project Principal Satisfaction with MLD Project

Satisfaction with MLD Project Services/Activities	Average Satisfaction*
Leadership coach working with me.	3.9
Leadership coach working with leadership team.	3.8
Leadership coach working with teachers.	3.7
STW Rubric criteria.	3.7
STW Rubric annual data report.	3.7
Professional development sessions at the national STW Conference.	3.5
Networking with other schools.	3.4
Professional development sessions in my state through the project.	3.3
Visits to other schools.	3.5
Hosting a mock STW visit at my school.	3.8
VAL-ED leadership assessment data report.	3.7
Teacher survey data report.	3.5

* 1=Not satisfied; 2=Somewhat satisfied; 3=Satisfied; 4=Very Satisfied.

> "The overall impact of the grant has been very positive. It has given us purpose, direction, and guidance to improve what we do for the overall growth of the whole student."
>
> - MLD Principal

When principals were asked about the barriers they encountered when implementing the MLD Project activities, there were no significant barriers reported. The two items that were indicated to be minor barriers included lack of time necessary for adequate planning/implementation and lack of clear expectations about what is involved in the project.

Implementation Findings

The purpose of the implementation study was to document the fidelity of implementation of the MLD Project at each treatment school, assessing whether the intervention was implemented as intended.

FIDELITY MATRIX

Using the MLD Project logic model, a measurement tool to track fidelity of implementation was developed (See Appendix A). The major project components were identified, as well as the key indicators to measure each major component. Then, for each indicator, an operational definition was developed, a data source identified, data collection schedule set, and an implementation score was determined so that an assessment of levels could be made (low, medium, high) for each indicator as well as across indicators for each major project component. Implementation fidelity was measured every year of the project (3 times), however, due to adjustments to project components after Year 1 of the project, only Year 2 and Year 3 implementation fidelity was used and analyzed.

The following data sources were used to assess fidelity of implementation: Coach's Log; state leadership team activity reports, vision statement documents, Self-Study Teacher Survey; STW Rubric; improvement plans and action plans; meeting minutes; phone call logs; Edmodo software records; sign-in and meeting attendance lists; and leadership team meeting notes.

Overall, each key project component was scored using multiple indicators. An implementation score was determined for each indicator. Some indicators were scored with a range of 0 to 1 where 0 is not completed and 1 is completed, while others had a range of 0 to 2 where 0 is low implementation, 1 is medium implementation, and 2 is high implementation. All indicators except one were measured at the school level. The one indicator not measured at the school level was measured at the teacher level instead. Once the indicator level scores were calculated, then for each project component, the individual indicator scores were scored at the key project component level to arrive at a component level implementation score. For the component level score, the approach was to total up the scores on each indicator and apply a component level score of 0 to 2 (0 is low implementation, 1 is medium implementation, 2 is high implementation).

Finally, for each component, a threshold was established for assessing fidelity of implementation across the sample (e.g., 85% percent of schools must have high implementation) in order to determine if fidelity of each component was met (Table 9).

IMPLEMENTATION FINDINGS

The following key project components were measured:

- Create a powerful vision for high performance using the STW criteria;
- Engage in assessment and planning to identify needs, develop goals, and implement an action plan;
- MLD leadership coach works with leadership team;
- Principal mentor works with principal;
- Network school works with school;
- School engages in a leadership professional learning community;
- School participates in networking experiences; and
- School receives regular and systematic professional development trainings.

Table 9. MLD Project Implementation Components and Thresholds

Key Programmatic Components	Data Source(s)	Implementation Score*	Threshold for Fidelity of Implementation for the Sample
Create a powerful vision for high performance using the STW criteria	Coach's Log, Online STW Rubric data collection system, Self-Study Teacher Survey	Low = 0-2 Medium = 3-7 High = 8-10	At least 70% of schools: Year 2 has medium or high implementation, Year 3 has high implementation
Engage in assessment and planning to identify needs, develop goals, and implement an action plan	Coach's Log	Low = 0-1 Medium = 2-7 High = 8-9	At least 70% of schools have high implementation
MLD leadership coach works with leadership team	State team activity reports, Coach's Log	Low = 0-2 Medium = 3-9 High = 10-11	At least 70% of schools have high implementation
Principal mentor works with principal	State team activity reports, Coach's Log	Low = 0 Medium = 2 High = 4	At least 70% of schools have high implementation
Network school works with school	State team activity reports	Low = 0 Medium = 2 High = 4	At least 70% of schools have medium or high implementation
School engages in a leadership professional learning community	State team activity reports	Low = 0-2 Medium = 3-5 High = 6-7	At least 80% of schools have high implementation
School participates in networking experiences	State team activity reports	Year 2: Low = 0, Medium = 1-2, High = 3-4	At least 70% of schools have high implementation

		Year 3: Low = 0, Medium = 1-2, High = 3	
School receives regular and systematic professional development trainings	State team activity reports	Low = 0-1 Medium = 2 High = 3	At least 70% of schools have high implementation

IMPLEMENTATION RESULTS

The research question guiding the implementation study: Is the MLD Project intervention implemented with fidelity at each treatment school? An examination of implementation fidelity across the entire sample for each of the eight key programmatic components highlighted stable implementation of the majority of the key components (Table 10). In Year 2, seven of the eight components were implemented with fidelity. The only component in Year 2 that was not met with fidelity was regular professional development trainings. In Year 3, six of the eight components were implemented with fidelity. In Year 3, the two components that were not met with fidelity included the leadership coach working with the leadership team, and the school participating in networking experiences. In focus group interviews with coaches, they commented that the majority of schools implemented all of the key project components with fidelity with only one or two schools having a low level of implementation of a few components.

Table 10. MLD Project Yearly Implementation with Fidelity by Component

Key Programmatic Components	Threshold Goal	Implemented with Fidelity for the Sample	
		Year 2	Year 3
		% of Sample Yes/No	% of Sample Yes/No
Create a powerful vision for high performance using the STW criteria	70%	100% Yes	92% Yes
Engage in assessment and planning to identify needs, develop goals, and implement an action plan	70%	83% Yes	75% Yes
MLD leadership coach works with leadership team	70%	75% Yes	58% No

Principal mentor works with principal	70%	92% Yes	75% Yes
Network school works with school	70%	75% Yes	83% Yes
School engages in a leadership professional learning community	80%	92% Yes	100% Yes
School participates in networking experiences	70%	100% Yes	50% No
School receives regular and systematic professional development trainings	70%	67% No	75% Yes

As mentioned previously, the MLD Project was implemented in multiple districts in Kentucky and Michigan by the National Forum affiliates (Kentucky Middle School Association-KMSA and Institute for Excellence in Education-IEE). Each state had varying requirements and priorities and the affiliates had established infrastructures for overseeing services to the schools. As a result of these differences, implementation of the networking and professional development activities of the project was approached in the same way that the National Forum uses the STW Rubric – by adapting to meet the needs of the schools involved and the realities of each setting. In Michigan, IEE used their central office location to host quarterly networking seminars for MLD Project schools. The principal and leadership team from each school attended the seminars where they participated in several key project components, including working with their leadership coach, working with their principal mentor, networking with other schools, and receiving professional development. It was an efficient way to implement the project activities while maximizing interactions among school personnel and the project team. In Kentucky, due longer distances between project schools which prohibited regular travel for networking meetings and professional development, project schools participated in project components at their individual school buildings with their leadership coach or other professional development trainer, and networking occurred when schools went on STW visits to other schools. All the schools across both states received the same services, however, the mechanism for delivering them varied. There were advantages to each delivery method. In Michigan, schools benefitted from more networking experiences. In Kentucky, schools received more individualized services at their buildings.

An examination of implementation results by school highlight several key findings related to the variability of implementation. The table below (Table 11) shows the level of implementation of each key project component by year and it also shows which components were determined by the project leaders to be weighted as most important to the achieve. The first observation is that overall, implementation was relatively high across all components at all schools. It is clear from

these data that four schools (School 4, School 5, School 8, and School 12) had slightly higher and more consistent overall implementation than the other schools. We hypothesize that implementation in these four schools may be higher because they had a higher level of readiness for making improvements at the start of the project and their district was highly involved and supportive of the project and its goals. Although overall implementation was moderately high, the four most highly implemented schools were examined to determine whether their level of implementation resulted in higher intermediate or long-term outcomes, and this was not the case. While these four schools had slightly higher implementation, it did not manifest in higher outcomes at this time. It may be the case that more time with the higher level of implementation in place is necessary for it to impact outcomes positively.

Table 11. MLD Project Yearly and Weighted Implementation by State and School

States	Schools	Key Programmatic Components Year 2/Year 3 Level of Implementation (High, Medium, Low)							
		Vision #1 Weight	Action	Coach	Mentor	Visits #2 Weight	PLC #1 Weight	NW #2 Weight	PD
KY	School 1	M/H	H/M	H/M	H/M	H/H	H/H	H/M	H/M
	School 2	M/H	H/H	H/H	H/H	H/H	H/H	H/M	M/H
	School 3	M/H	H/H	M/M	L/H	M/H	H/H	H/M	L/H
	School 4	H/H	H/H	H/H	H/H	H/H	H/H	H/H	M/H
	School 5	H/H	H/H	H/H	H/H	H/H	H/H	H/M	H/H
	School 6	M/H	H/H	H/H	H/H	M/H	H/H	H/H	M/H
MI	School 7	H/M	H/M	M/M	H/H	H/H	H/H	H/M	H/L
	School 8	H/H	H/H	H/H	H/H	H/H	H/H	H/H	H/H
	School 9	M/H	M/M	M/M	H/H	M/M	M/H	H/H	H/H
	School 10	M/H	H/H	H/M	H/M	H/H	H/H	H/H	H/M
	School 11	H/H	M/H	H/H	H/M	H/M	H/H	H/M	H/H
	School 12	H/H	H/H	H/H	H/H	H/H	H/H	H/H	H/H
		7M 17H	5M 19H	8M 16H	1L 3M 20H	5M 19H	1M 23H	6M 18H	2L 5M 17H
	Highest Implementation (all high or all but 1 high)								
	Medium Implementation (3 or 4 medium)								
	Lowest Implementation (5, 6, or 8 medium; 2 low)								

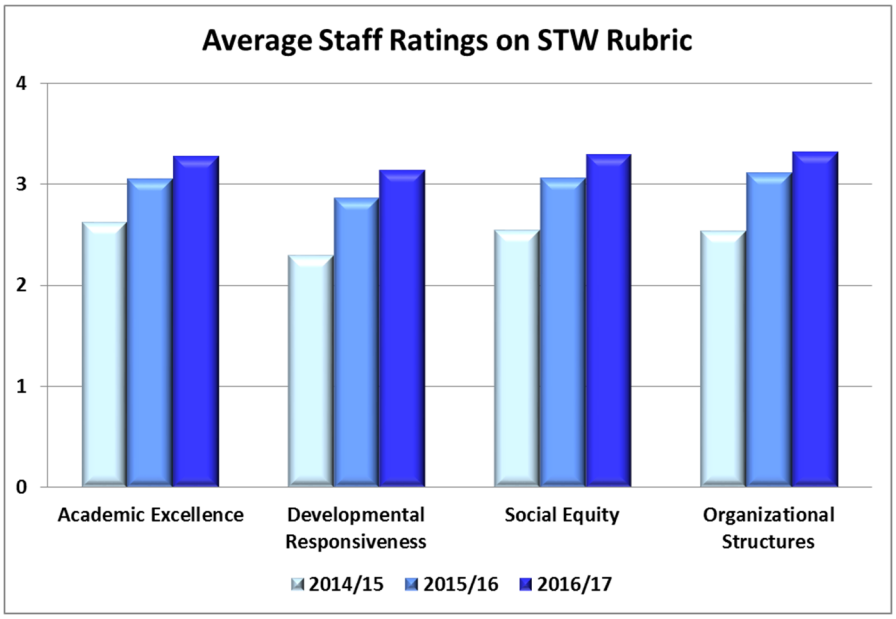
Intermediate Outcome Findings

STW RUBRIC

> "Our intentionality over the past two years in working with the STW Rubric as well as being proactive instead of reactive, has helped us to all get on the same page and remain true to our vision. We're very focused now."
> - MLD Principal

An analysis of the yearly aggregated STW Rubric data using independent sample t-tests showed that there was **statistically significant improvement in staff ratings of their implementation of each of the four key STW Rubric criteria from Year 1 to Year 3** of the MLD Project (Figure 2). Overall, staff reported significantly higher implementation of the STW criteria ($p \leq .001$) from Year 1 to Year 3.

Figure 2. MLD Project STW Rubric Results by Year



4=High quality; 3=Good quality; 2=Fair quality; 1=Poor quality p≤.001

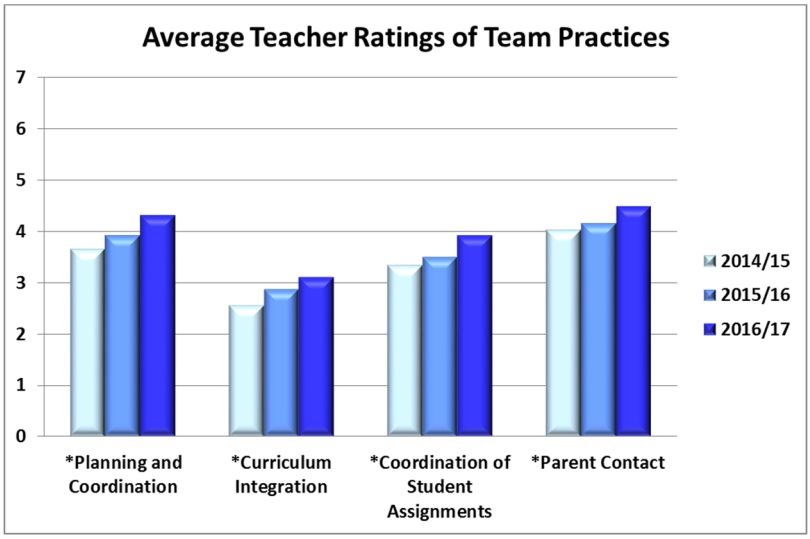
SELF-STUDY TEACHER SURVEY

A series of independent samples t-tests were run to determine if there were differences in the Self-Study Teacher Survey constructs from Year 1 to Year 3 of the MLD Project. Teachers reported **statistically significant improvements in the following constructs related to collaboration, shared leadership, and classroom instructional practices.**

Team Practices increased – Teachers reported engaging more frequently in planning and coordination activities, curriculum integration, coordination of student assignments and

assessments, and contract with parents (Figure 3). These improvements were statistically higher in Year 3 than in Year 1 ($p \leq .05$).

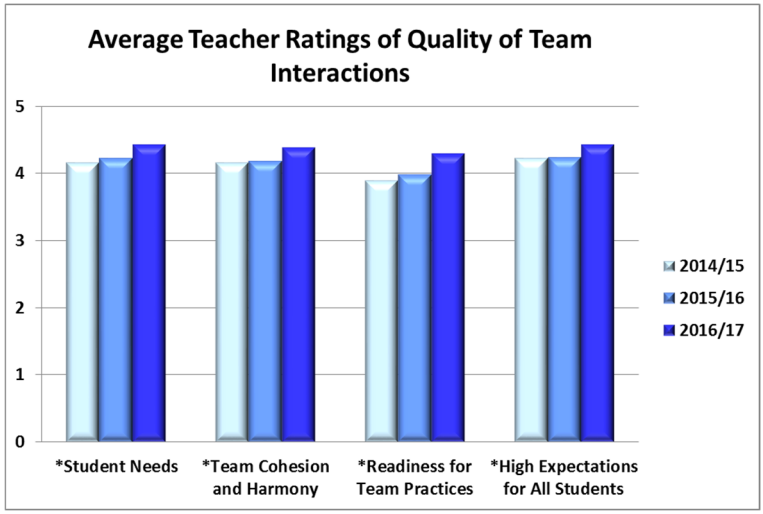
Figure 3. MLD Project School Improvement Self-Study Teacher Survey: Team Practices by Year



7=Daily; 6=Weekly; 5=Monthly; 4=Quarterly, 3=Several times a year, 2=Once a year; 1=Never p≤.05

Quality of Team Interactions Improved – Cohesion and harmony among teachers working together improved, teachers felt more prepared to work together in a collaborative way, and teachers agreed that they not only address student needs but have consistently high expectations for students (Figure 4). These improvements were statistically higher in Year 3 ($p \leq .05$).

Figure 4. MLD Project School Improvement Self-Study Teacher Survey: Quality of Team Interactions by Year

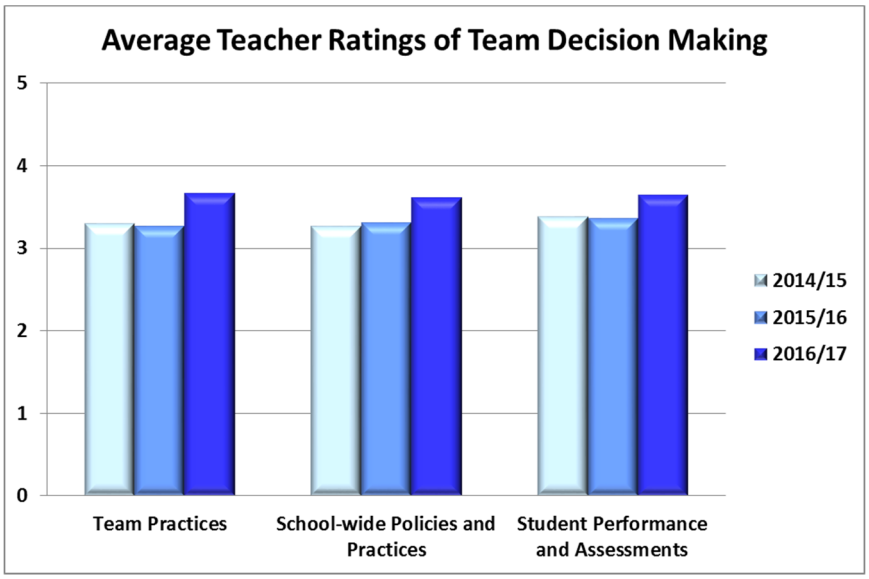


5=Strongly agree; 4=Agree; 3=Neither agree nor disagree; 2=Disagree, 1=Strongly disagree p≤.05

INTERMEDIATE OUTCOME FINDINGS

Team Decision Making Increased – Teachers reported increased decision making opportunities among their teams regarding team practices, school-wide policies and practices, and student performance and assessments (Figure 5). These improvements were statistically higher in Year 3 ($p \leq .05$).

Figure 5. MLD Project School Improvement Self-Study Teacher Survey: Team Decision Making by Year



5=Very much; 4=Much; 3=Moderate amount; 2=Little; 1=Very little p≤.05

Teacher Efficacy and Collective Teacher Efficacy Improved – Teachers reported a statistically significant ($p \geq .05$) improvement in their individual, as well as collective, commitment to the success of their students by Year 3 of the grant. In other words, teachers reported stronger beliefs that they and their colleagues have what it takes to get students to learn, are able to get through to difficult students, are confident they will be able to motivate their students, and believe that if a student does not learn something the first time, teachers will try another way.

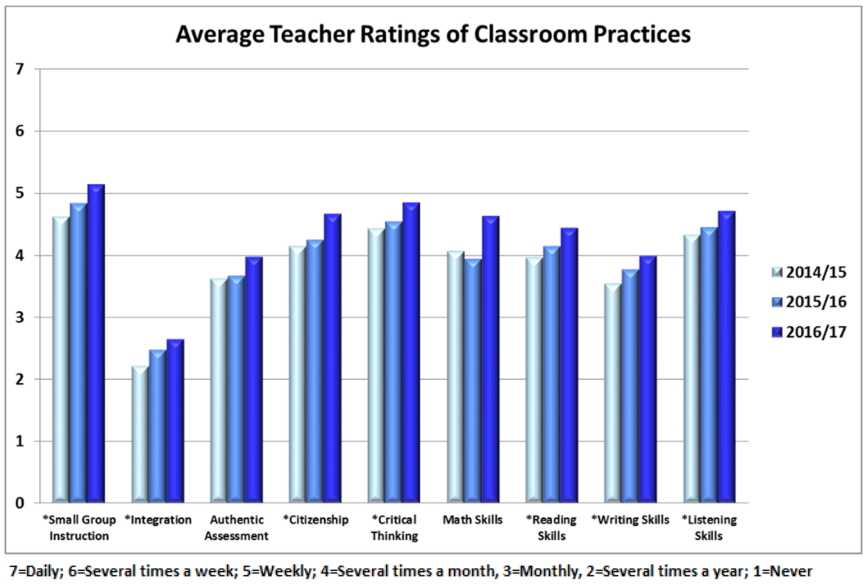
Teachers report **significant increases** in the use of "best" **middle-grades instructional practices** (Figure 6). Teachers report engaging more frequently in the following instructional practices:

- Small group active instruction
- Integration and interdisciplinary practices

- Citizenship and social competence instruction
- Critical thinking practices
- Reading skill practices
- Writing skill practices
- Listening and verbal skill practices

> "The STW Rubric helps with ownership of our improvement goals. It helps with the impetus for the instructional shifts that need to happen. That focus for the principal then being able to be proactive, not trying to always catch up, but leading. Truly leading."
>
> - MLD Leadership Coach

Figure 6. MLD Project School Improvement Self-Study Teacher Survey: Classroom Practices by Year



COACH'S LOG

The 12 MLD Project schools in Kentucky and Michigan received coaching services from a leadership coach from fall 2014 through spring 2017. The leadership coaches completed an online Coach's Log for any key or substantive visit, conference call, email communication, or webinar with their school. Overall, coaching services were a very intensive and ongoing aspect of the MLD Project. Throughout the three-year project, leadership coaches made 715 contacts with project schools consisting of 1,103 different activities for more than 3,600 hours (Table 12). This amounted to an average of 59 contacts per project school.

Table 12. MLD Project Coaching Contacts and Activities by Year

Annual Contacts and Activities	Coaching Services			
	Year 1	Year 2	Year 3	Total
Number of schools served	12	12	12	12
Number of contacts	191	268	256	715
Number of hours	853	1,475	1,359	3,687
Number of activities	320	346	437	1,103
Average contacts per school	16	22	21	59
Average hours per contact	4.5	5.5	5.3	5.1

Although there were some key elements of coaching that were ongoing throughout all three years of the MLD Project (i.e., providing generalized coaching services; conducting walkthroughs and observations; and working with schools to use data effectively for decision making), there were other coaching services that evolved in response to the progress and needs of the school (Table 13).

Table 13. MLD Project Key Coaching Elements by Year & Descriptions

<table>
<tr><th colspan="3">Key Coaching Elements</th></tr>
<tr><td>Year 1 (2014/15)</td><td>Year 2 (2015/16)</td><td>Year 3 (2016/17)</td></tr>
<tr><td>Key Elements in Year 1
• Developing trusting relationships; and
• Learning about the school</td><td></td><td></td></tr>
<tr><td colspan="3">Key Elements in Year 1, Year 2, and Year 3
• Providing generalized coaching services (e.g., mentoring, providing resources);
• Conducting walkthroughs and observations; and
• Working with schools to use data effectively for decision making</td></tr>
<tr><td rowspan="2"></td><td colspan="2">Key Elements in Year 2 and Year 3
• Capitalizing on networking opportunities (e.g., state workshops, annual STW Conference);
• Utilizing STW visits to drive change;
• Utilizing mock STW visits to improve teaching practices;
• Assessing school readiness to apply for STW designation; and
• Focusing on informed and fluid RTI Tiered services</td></tr>
<tr><td></td><td>Key Element in Year 3
• Planning for sustainability; and
• Applying for STW designation</td></tr>
</table>

INTERMEDIATE OUTCOME FINDINGS

In Year 1, coaches focused their efforts on developing trusting relationships, as well as, learning about their schools in order to provide more responsive coaching. Starting in Year 2 and continuing into Year 3, coaches reported capitalizing on network opportunities; utilizing STW visits to drive change; utilizing mock STW visits to improve teaching practices; assessing their readiness to apply for STW designation; and focusing on informed and fluid RTI Tiered services. In Year 3, coaches focused on planning for sustainability with the end of the grant approaching. Finally, although schools explored their "readiness" in Year 2, it was not until Year 3 that schools were applying for STW designation with their coach's assistance. The key coaching elements are further described in the section below.

Descriptions of Key Coaching Elements

- **Developing trusting relationships** with the principal, leadership team, and teachers;
- **Learning about the school** in order to prioritize needs and provide responsive coaching;
- **Providing generalized coaching services,** such as coaching conversations and mentoring; participating in team *(e.g., administration, leadership, PLCs)* meetings; assisting with state and district requirements; helping with the creation of the master schedule; advising on the recruitment and hiring of staff; attending school events; and providing resources and professional development to the principal, teachers, and school;
- **Conducting walkthroughs and observations** *(e.g. classrooms, PLCs, school)* utilizing protocols or observation tools *(e.g., STW Rubric, school district instrument)* with a focus *(e.g., determining professional development needs and levels of engagement, assessing progress on the STW Rubric and best instructional practices)*; debriefing/calibrating with an administrator or the leadership team; providing feedback and strategies to teachers; shadowing principals and assistant principals to observe their interactions with staff and students; and facilitating peer observations;
- **Working with schools to use data effectively for decision making** by utilizing multiple data sources *(e.g., STW Rubric, Teacher Survey, VAL-ED Assessment, state achievement scores, mock visit, walkthrough/classroom observations, discipline and referral data)*, disaggregations, and longitudinal data for: facilitating the principal's reflection on their practice and the development of a professional growth plan; utilizing formative assessment data to impact teacher instructional practices and address student learning styles; establishing and adjusting RTI tiered services; monitoring early indicators to better support students; informing (and adjusting) student discipline policies and programs; aligning state and district expectations, initiatives, and action plans with the school's improvement plan; and celebrating successes;

- **Capitalizing on networking opportunities** (i.e., STW Conference, networking meetings, STW site visit training) by sharing successes and resources from other coaches, project schools, and STW schools; debriefing after sessions and discussing how to implement best practices and other takeaways at their school; building capacity through opportunities to practice skills (*e.g., collaboration skills*) during networking meetings; and learning more about the STW process through participation in site visit trainings;
- **Utilizing STW visits to drive change** by determining the school visit focus then selecting the relevant school and teachers to visit; discussing observations and questions with the visited school teachers and administration; collaboratively debriefing and discussing takeaways; communicating these observed best practices and programs to the rest of the school; and adapting/implementing practices and programs to meet the needs of their school;
- **Utilizing mock STW visits to improve teaching practices** by monitoring instruction more deeply with the STW Rubric in preparation of the visit and the visiting team's expectations; debriefing with the visiting team's constructive feedback; validating teaching practices and confirming administrator assessments; and discussing "applause" and "concerns" to direct further instructional changes;
- **Assessing school readiness to apply for STW designation** by reviewing the application, explaining the process, and identifying the necessary data to prepare; gathering information at the STW Conference and through participation in STW visits and a STW site visit training; meeting with teams and individual stakeholders *(e.g., principal, administrative team, leadership team, teacher(s), central office)* to assess their perceptions, concerns, and support of the application process and timeline; and either deciding to apply during Year 3, or making plans to apply in the following year;
- **Focusing on informed and fluid RTI tiered services** by utilizing data at the beginning, middle, and end of the year to inform student placement; analyzing areas of weakness to develop lessons for mastery and novice reduction in math, ELA, and ELL for individual students; re-teaching standards with lowest scores; and looking at teacher evaluation data and certifications to refine both the master schedule focusing on RTI time, as well as needed teacher professional development;
- **Planning for sustainability** by introducing the sustainability protocol and matrix (i.e., assigning initiatives to quadrants, reflecting on priorities, and creating an action plan); planning to support ongoing practices for next year; understanding the "why" instead of just focusing on the "what" and "how"; "ramping up" distributive leadership; "institutionalizing" changes to "survive" staffing changes—especially the principal and superintendent; developing an orientation for new faculty on the school's culture and way of doing things;

taking on roles and responsibilities that the leadership coach was providing; utilizing the STW Rubric for reflecting on strengths, challenges, and next steps in sustaining best practices; and focusing on mediating questions from mock visits and pre-visits to refine practices and consistency throughout school;

- **Applying for STW designation** by assessing school readiness to apply; engaging with the administrative team, leadership team and faculty in collaboratively gathering evidence to assess and align their school with the STW Rubric, as well as make any needed changes to strengthen their school's use of the STW criteria; examining and disaggregating achievement data to discuss how to make meaningful changes to reach gap students; explaining the application process and STW designation to parents, school board, and the community; reflecting on multiple iterations of the application before submitting an application; visiting individual classrooms to make specific recommendations to teachers; conducting school walkthroughs to measure school readiness for the designation visit; and hosting the visiting designation team.

  The journey continued after their successful STW designation with much celebrating by staff, students, and stakeholders *(e.g., school district, school board, politicians)* and the presentation of the STW banner—or multiple presentations of the STW banner with different audiences. Afterwards schools reported: being ready for new initiatives; reviewing exit report notes; having open and frank conversations about growth areas in order to determine what to highlight in their STW School presentation at the annual conference; presenting to the school board; hosting their first non-project school for a visit; and celebrating again at the STW Conference.

FOCUS GROUPS

Qualitative data analysis of the principal and leadership coach focus groups that occurred after each year of the MLD Project provided formative data on the quality and fidelity of implementation, as well as the change process. The meaningful content of the qualitative data were examined using an inductive approach. Recurrent themes were identified with a focus on whether the content represented group-shared ideas. Through this content analysis, the data were summarized, with the highlights below:

Year 1 Findings:

- Coach's successful process to build trust, credibility, and rapport with principal and staff to facilitate responsive coaching;

- Coach's ongoing process to mentor and model a collaborative, distributive leadership style for the principal;
- Development of teacher leaders through the creation of leadership teams;
- Utilization of project data (i.e., STW Rubric, Self-Study Teacher Survey, VAL-ED) in a purposeful manner to secure buy-in from staff for implementation of the MLD Project;
- Identification of goals and challenges for implementing the project in Year 2; and
- Coach's assessment that the implementation of the MLD Project had positively impacted school culture.

Year 2 Findings:

- School's focus on developmental responsiveness and student needs in Year 2 to improve school climate and culture—most frequently through the adoption of an advisory program;
- Improvements in collaborative leadership owing to principals creating a supportive risk-taking environment and intentionally growing teacher leaders with increased decision-making power and autonomy;
- Positive impact of developing principal and teacher leadership through professional development and networking opportunities through the MLD Project;
- Improving the principal's leadership skills through reflective conversations with the coach on the VAL-ED data to identify strengths and weaknesses in order to develop a professional growth plan with targeted goals;
- Process of developing teachers as instructional leaders, and building principal capacity to support teachers in the classroom; and
- Plans to build capacity for sustainability beyond the grant.

Year 3 Findings:

- Key ingredients for replicating success in transforming project schools into nationally recognized STW Schools;
- Descriptions of unexpected impacts of participation in the MLD Project that positively impacted project schools;
- Principal assessment that their school's collaborative leadership has improved as a result of participation in the MLD Project;
- Importance of building school culture focused on the holistic child with a student-centered environment, which principals and coaches believed improved academics;
- Principal assessment that developing teacher leaders for a collaborative, distributive leadership model is the key to building capacity for sustainability beyond the grant or principal turnover; and
- Recommendations for the implementing the MLD Project in future work.

Final Outcome Findings

PRINCIPAL EFFECTIVENESS

The final outcome of principal effectiveness was measured with the VAL-ED Leadership Assessment. VAL-ED provides scores for six core components of leadership: high standards for student learning; rigorous curriculum; quality instruction; culture of learning and professional behavior; connections to external communities; and performance accountability. VAL-ED further breaks down six key leadership processes (planning, implementing, supporting, advocating, communicating, and monitoring) within each core component to provide an in-depth picture of each principal's leadership. At the project level, principals improved for each core component and key process from Year 1 to Year 3 of the MLD Project, with principals demonstrating the greatest growth in the areas of culture of learning and professional behavior, quality instruction, performance accountability, and high standards for student learning (Table 14).

> "My superintendent asked, 'You sure you want to have your staff take the VAL-ED?' Yeah, I do. I mean this is my community. These are the people I work with, and if they see me doing something that could become better, why not become better? So that's check your ego at the door and just do what you need to do."
>
> - MLD Principal

Leadership skills and behaviors also improved as measured by the principal's overall effectiveness score, which is indicated by both a mean score and a percentile score. The mean score is the average of all respondent groups (i.e., principal, teachers, supervisor) on all key processes and core components to determine the individual principal's performance level. VAL-ED assigns performance levels (i.e., distinguished, proficient, basic, below basic) according to set ranges of mean scores that do not change based on the national sample. The number of schools with proficient or distinguished principals grew from two schools in Year 1 to nine schools by the end of Year 3 (Table 15). The percentile score is the principal's mean score in comparison to a national sample of principal means. Based on the percentile score, twelve principals grew as leaders when compared to the national sample.

Table 14. MLD Project VAL-ED Principal Effectiveness Scores by Year

Average Principal Effectiveness Scores on Leadership Core Components and Key Processes			
	Year 1 (2014/15)	Year 2 (2015/16)	Year 3 (2016/17)
High Standards for Student Learning			
Planning	3.44	3.53	3.72
Implementing	3.36	3.65	3.84
Supporting	3.48	3.66	3.71
Advocating	3.21	3.48	3.63
Communicating	3.23	3.78	3.52
Monitoring	3.48	3.75	3.88
Rigorous Curriculum			
Planning	3.14	3.32	3.53
Implementing	3.04	3.26	3.54
Supporting	3.67	3.74	3.93
Advocating	3.17	3.43	3.70
Communicating	3.41	3.71	3.70
Monitoring	3.10	3.52	3.52
Quality Instruction			
Planning	3.37	3.78	3.85
Implementing	3.49	3.47	3.83
Supporting	3.81	3.80	4.14
Advocating	3.39	3.46	3.73
Communicating	3.36	3.62	3.87
Monitoring	3.49	3.60	3.88
Culture of Learning and Professional Behavior			
Planning	3.59	3.58	3.87
Implementing	3.66	3.85	4.02
Supporting	3.69	3.85	4.01
Advocating	3.43	3.63	3.77
Communicating	3.28	3.74	3.76
Monitoring	3.14	3.78	3.70
Connections to External Communities			
Planning	2.98	3.31	3.39
Implementing	2.92	3.53	3.24
Supporting	2.93	3.59	3.56
Advocating	2.80	3.49	3.36
Communicating	3.33	3.55	3.74
Monitoring	3.03	3.33	3.52
Performance Accountability			
Planning	3.23	3.67	3.66
Implementing	3.28	3.41	3.68

Supporting	3.43	3.42	3.79
Advocating	3.34	3.37	3.66
Communicating	3.35	3.28	3.50
Monitoring	3.29	3.43	3.63
VAL-ED Performance Levels			
Below Basic (1.00-3.28)	Basic (3.29-3.59)	Proficient (3.60-3.99)	Distinguished (4.00-5.00)

Table 15. MLD Project VAL-ED Principal Overall Effectiveness Scores and Performance Levels by Year and School

Principal Overall Effectiveness Scores and Performance Levels

School *	Mean Score				Percentile Score			
	Year 1	Year 2	Year 3	Change	Year 1	Year 2	Year 3	Change
School 1	2.38	3.97	4.25	1.87	0.4%	83.4%	95.3%	94.9%
School 2	3.78	4.36	4.17	0.39	69.3%	97.6%	92.6%	23.3%
School 3	3.6	3.56	4.07	0.47	49.2%	45.6%	88.4%	39.2%
School 4	3.07	3.29	3.82	0.75	6.0%	17.0%	73.0%	67.0%
School 5	3.26	3.59	3.79	0.53	14.1%	47.8%	70.2%	56.1%
School 6	3.2	3.19	3.68	0.48	12.3%	11.4%	59.3%	47.0%
School 7	3.5	3.94	3.66	0.16	36.0%	82.1%	57.0%	21.0%
School 8	3.52	3.48	3.64	0.12	39.2%	34.6%	55.1%	15.9%
School 9	3.49	3.49	3.63	0.14	35.0%	36.0%	54.0%	19.0%
School 10	3.32	3.52	3.45	0.13	18.2%	38.7%	30.1%	11.9%
School 11	3.16	3.27	3.41	0.25	10.0%	14.5%	26.4%	16.4%
School 12	3.29	3.51	3.38	0.09	16.8%	37.4%	22.8%	6.0%
School 13	3.29	3.00	3.14	-0.15	17.3%	5.0%	8.6%	-8.7%

*Two schools had principal turnover during the 3-year MLD Project.

VAL-ED Performance Level:
Below Basic (1.00-3.28)
Basic (3.29-3.59)
Proficient (3.60-3.99)
Distinguised (4.00-5.00)

FINAL OUTCOME FINDINGS

STUDENT ACHIEVEMENT TEST SCORES

The final outcome of student achievement was measured with the state achievement tests: the Kentucky Performance Rating for Educational Progress (K-PREP); and the Michigan Student Test of Educational Progress (M-STEP) test. Each test was administered annually in 2015, 2016, and 2017 at all project and comparison schools.

Overall school achievement

Overall, in the state of Kentucky, student achievement on the K-PREP test in 2017 increased slightly at the middle school level with achievement gaps between different groups of students persisting in many areas. The percentage of middle school students performing at the Proficient/Distinguished levels on the KPREP test increased in reading, social studies, and language mechanics. A comparison of the MLD project schools to the state at the 8th grade level showed that two schools had larger reading growth than the state and two schools had larger math growth than the state (Table 16). See Appendix B for the full table containing all grade levels.

Table 16. MLD Project K-PREP Achievement for Overall State and 8th Grade Project Schools by Year

Reading K-PREP Results 2015-2017	**% of Proficient or Advanced**			**Growth 2015 to 2017**
	2015	**2016**	**2017**	
8th Grade				
Overall State	54.1%	53.6%	57.1%	3.0%
School 1	42.5%	46.0%	45.3%	2.8%
School 2	55.9%	44.0%	50.9%	-5.0%
School 3	43.9%	43.4%	51.5%	7.6%
School 4	57.6%	48.0%	47.3%	-10.3%
School 5	62.9%	58.4%	66.5%	3.6%
School 6	46.3%	50.9%	47.9%	1.6%
Mathematics K-PREP Results 2015-2017	**% of Proficient or Advanced**			**Growth 2015 to 2017**
	2015	**2016**	**2017**	
8th Grade				
Overall State	44.2%	45.5%	48.7%	4.5%

School 1	46.8%	42.9%	40.7%	-6.1%
School 2	37.4%	44.0%	34.0%	-3.4%
School 3	19.7%	24.1%	27.6%	7.9%
School 4	33.7%	31.5%	35.7%	2.0%
School 5	50.9%	49.8%	59.9%	9.0%
School 6	36.8%	48.5%	38.7%	1.9%
	Higher than the overall state			

Overall in the state of Michigan, student achievement on the M-STEP achievement test in 2017 increased for mathematics in the middle grades of 6, 7 and 8. The Michigan Department of Education, however, noted "the English language arts scores are disappointing" in that ELA scores decreased in 2017 in grades except grade 5. A comparison of the MLD project schools to the state at the 8th grade (note that school level results are not available) showed that five schools had larger ELA growth than the state and one school had larger mathematics growth than the state (Table 17). See Appendix B for the complete table containing all grade levels.

Table 17. MLD Project M-STEP Achievement for Overall State and 8th Grade Project Schools by Year

ELA M-STEP Results 2015-2017	**% of Proficient or Advanced**			**Growth 2015 to 2017**
	2015	**2016**	**2017**	
8th Grade				
Overall State	47.6	48.8	48.0	0.4%
School 7	50.4	46.7	51.4	1.0%
School 8	51.9	53.2	58.4	6.5%
School 9	33.8	31.7	28.6	-5.2%
School 10	41.5	57.1	47.4	5.9%
School 11	43.9	46.3	50.0	6.1%
School 12	30.8	25.4	32.2	1.4%
Mathematics M-STEP Results 2015-2017	**% of Proficient or Advanced**			**Growth 2015 to 2017**
	2015	**2016**	**2017**	

8th Grade				
Overall State	32.2	32.7	33.5	1.3%
School 7	37.0	26.2	29.9	-7.1%
School 8	34.6	40.6	50.8	16.2%
School 9	6.8	8.4	7.7	0.9%
School 10	35.4	39.4	35.1	-0.3%
School 11	38.6	29.6	28.1	-10.5%
School 12	23.1	9.9	13.6	-9.5%
Higher than the overall state				

Impact study final analytic sample

As described in the Evaluation Design section, the impact study tracked a cohort of project and comparison students starting in 6th grade in the 2014/15 school year over three years by matching individual students over time. The final analytic sample for ELA/reading included 1,300 intervention students and 4,412 comparison students. The final analytic sample for math included 1,304 intervention students and 4,433 comparison students. Students were from a total of 50 schools (12 intervention and 38 comparison). Demographics of the sample are in Table 18.

Table 18. MLD Project Demographics of Final Impact Study Analytic Sample

		Intervention Group		Comparison Group	
		Count	Percent (%)	Count	Percent (%)
Gender	Male	668	51.2%	2,186	49.3%
	Female	636	48.8%	2,247	50.7%
Free/Reduced Lunch	Yes	715	54.8%	2,639	59.5%
	No	589	45.2%	1,794	40.5%
Race	White	1,077	82.6%	3,662	82.6%
	African American	69	5.3%	240	5.4%
	Hispanic	103	7.9%	288	6.5%
	Other	55	4.2%	243	5.5%
Special Education Status	No	1,197	91.8%	3,963	89.4%
	Yes	107	8.2%	470	10.6%

English Language Learner Status	**No**	1,279	98.1%	4,364	98.4%
	Yes	25	1.9%	69	1.6%

Impact study baseline equivalence

Baseline equivalence was conducted using the final analytic sample. For confirmatory research question one, 5th grade math achievement scores were used, and for confirmatory research question two, 5th grade ELA/reading achievement scores, were used to examine the equivalence between treatment and comparison schools. A 2-level model, where students were nested within schools, was performed to estimate the treatment minus comparison difference as done in the impact analysis without including any other covariates into the model. The structure of the model is shown below:

Level 1:

$Y_{ij} = \beta_{0j} + \varepsilon_{ij}$

Level 2:

$\beta_{0j} = \Upsilon_{00} + \Upsilon_{01}(Trt) + \mu_{0j}$

Where

Y_{ij} =5th grade measure (ELA/reading or math) of the i^{th} student in the j^{th} school.

Trt = 1 if in treatment schools, =0 if in comparison schools

ε_{ij} = residual error, $\varepsilon_{ij} \sim N(0, \sigma^2)$ and $\mu_{0j} \sim N(0, \tau^2)$

Fitting the above model to the data would produce an estimate of Υ_{01}, which represents the estimated mean difference in the pre-test measures.

For unadjusted means and standard deviations, descriptive statistics were computed for the treatment and comparison groups for both ELA/reading and math. Then, the standardized difference was calculated as the model adjusted mean difference divided by the pooled standard deviation using this formula –

$$S_{pooled} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 + n_2 - 2)}}$$

Where n_1 and n_2 are the student sample sizes, and S_1 and S_2 are the unadjusted standard deviations for the intervention and comparison group, respectively.

The standardized difference was then multiplied by a factor of $\omega = [1 - 3/(4N - 9)]$, with N being the total sample size. Because of the large sample sizes, ω gets close to one and as a result this correction does not make much difference of the original standardized difference.

There were no significant differences in baseline means (Table 19). In other words, the intervention students and the comparison students were equivalent on achievement test scores prior to the start of the intervention. Therefore, any differences in achievement between the groups after the intervention can be considered a result of the intervention and not because the groups began at different levels of achievement. Although the final analytic sample did not show any baseline difference, since the standard deviation units exceeded .05, adjustments were made in the analyses at both the student and school levels.

Table 19. MLD Project Baseline Balance Testing Results of Student Achievement Using Final Analytic Sample

Outcome Domain	Intervention Group			Comparison Group			Model Adjusted Mean Difference	Standardized Difference Effect Size w/ Hedges' Correction	P-value
	Unadjusted Mean	Unadjusted Standard Deviation	Student Sample Size	Model-adjusted Mean	Unadjusted Standard Deviation	Student Sample Size			
ELA/Reading - 5th grade	-0.0842	0.9458	1,300	-0.0089	0.9599	4,412	-0.0648	-0.068	0.223
Math - 5th grade	-0.0877	0.8997	1,304	-0.0117	0.9316	4,433	-0.0571	-0.062	0.403

Impact study statistical models

The final analytic sample, evaluating the impact of the intervention, consists of 8th grade achievement scores after following the cohort of 6th graders who continued to belong to the same treatment schools as stated in Evaluation Design. Any students with missing pre-test achievement scores were excluded from the analytic sample. In addition, any students who moved during the project years were not included in the analysis. For confirmatory research question 1, 8th grade state standardized math achievement scores were the primary outcome. For confirmatory research question 2, 8th grade state standardized ELA/reading achievement scores were the primary outcome. The impacts of the intervention were tested on the combined cohorts of Kentucky and Michigan.

From the Unconditional Means Models, an ICC of 4.3% for ELA/reading and an ICC of 8.2% for math suggested the use of Hierarchical Liner Models (HLMs). A series of 2-level models, where

students were nested within schools, were executed to analyze the intervention effect on student achievement.

Let i refer to students and j refer to schools. Let Y_{ij} be 8th grade achievement (math or ELA/reading) for student i in school j. The model used to estimate the impact for a particular contrast has the two-level structure of students nested in schools, as shown below:

Level 1:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(Y_{pre.ij}) + \sum_{k=1}^{K} \beta_{k+1}(x_{ijk}) + \varepsilon_{ij}$$

Level 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(Trt) + \sum \gamma_{0j}(w_{ij}) + \mu_{0j}$$

Where

Y_{ij} = 8th grade achievement scores (ELA/reading or math) of the i^{th} student in the j^{th} school.

$Y_{pre.ij}$ = 5th grade achievement scores (ELA/reading or math) of the i^{th} student in the j^{th} school.

x_k represents k=1, 2, … , K student level covariates included in the model.

w_j represents j=1, 2, … , J school level covariates included in the model.

Trt = 1 if in treatment schools, =0 if in comparison schools

ε_{ij} = residual error, $\varepsilon_{ij} \sim N(0, \sigma^2)$ and $\mu_{0j} \sim N(0, \tau^2)$ and independent of ε_{ij}.

Fitting the above model to the data would produce an estimate $\hat{\gamma}_{01}$ that represents the estimated impact of treatment.

The following student and school-level covariates were used in the analyses:

Student-level covariates – 5th grade pre-test achievement scores, race/ethnicity (white vs. not-white, African American vs. not African American, Hispanic vs. non-Hispanic), gender, LEP status, free/reduced lunch status, and special education status; and

School-level covariates – free/reduced lunch, enrollment, NCLB AYP status, and school demographics.

FINAL OUTCOME FINDINGS

Other than the baseline 5th grade achievement scores, the other covariates did not change much of student and school level variances. However, all possible covariates were included as part of adjusting for any baseline differences with the goal of improving the accuracy of the coefficients and ensuring alignment with the covariates used in the propensity scores while matching comparison schools. STATA xtmixed was used to execute the models. In addition, a correlation of 0.69 to 0.71 was established between baseline 5th grade achievement scores and the 8th grade achievement outcome for both ELA/reading and math.

The effect size was calculated using Hedges's *g* ES index. It is defined as the standardized mean difference, which is the difference between the mean outcome of the treatment group and the mean outcome of the comparison group, divided by the pooled within-group standard deviation (SD) on 8th grade achievement scores. Since the achievement scores were at the student-level, the student-level standard deviation was used.

The final models, adjusting for all possible covariates, suggest no significant intervention effect on either ELA/reading scores or math scores (Table 20). In other words, the students that received the MLD Project intervention performed the same on both the ELA/reading test (Appendix C) and the math test (Appendix D) as the comparison students after three years. Neither the p-values nor effect sizes suggest a significant intervention effect on achievement.

Table 20. MLD Project Impact Study Results

Outcome Domain	Intervention Group				Comparison Group				Impact Estimate (Model Adjusted Mean Difference)	Standardized Effect Size w/ Hedges' Correction	Impact Standard Error	P-value
	Cluster Sample Size	Student Sample Size	Unadjusted		Cluster Sample Size	Student Sample Size	Unadjusted					
			Mean	Standard Deviation			Mean	Standard Deviation				
ELA/ Reading	12	1,300	-0.0178	0.936	38	4,412	-0.0223	0.955	0.0431	0.045	0.0551	.434
Math	12	1,304	-0.0898	0.877	38	4,433	-0.0169	0.935	-0.0005	-0.0005	0.0816	.995

Even though the impact analyses did not find an overall intervention effect on either ELA/reading scores or math scores, it should not be interpreted to mean that the MLD Project intervention was not effective, but that project schools may need additional time to fully implement and refine the model with middle grades students to advance achievement scores. We can further hypothesize that the positive changes noted in prior sections of this report on school organization and climate factors such as leadership, culture, climate, and instructional practices are a positive precursor to improvements in achievement since achievement does not improve without corresponding improvements in the teaching and learning environment.

Unexpected Findings/Spillover Effects

While the MLD Project logic model (Figure 1) clearly illustrates the anticipated intermediate and long-term outcomes expected from the implementation of the eight programmatic components, in practice there were additional, unexpected outcomes as a result of the spillover effect that benefitted the project schools, state leadership teams, and the project leadership team. This section highlights the most significant findings.

STW DESIGNATIONS

Although the focus of the MLD Project was to lay a foundation for a school's journey towards a National Forum STW designation, it was not an expectation before the end of the project. That 7 of the 12 project schools applied for and successfully received a STW designation before the end of the project (and another 3 schools during the no-cost extension), is a major triumph of the project and a tribute to the hard work of all of the stakeholders (i.e., administrators, faculty, students, and coaches) involved. Coaches believed that teachers and administrators had a "growth mindset" and an undefinable "readiness factor" to improve their schools. In addition to the support of the school district the following 5 areas of focus from the MLD Project were deemed essential to their success: 1) support from the leadership coach; 2) STW Rubric to focus school improvement; 3) transformational power of STW visits; 4) adoption of a distributive, collaborative leadership model; and 5) creation of a student-centered learning environment.

MOCK STW VISITS

Although the mock STW visits were not a planned program component, they were added in Years 2 and 3 of the MLD Project. The objective was to support project schools on the pathway towards a STW School designation after the grant ended. With this end goal in mind, the project team replicated key elements of the STW application process and school visit to provide project schools with a rich simulation, while supporting them with the expertise of their leadership coach and a national team of STW-trained middle-grades educators (see Lessons Learned section for more information about the team and process). The STW process and feedback provided the majority of schools with the confidence and tools needed to apply for STW designation. However, an added benefit was that principals, school leadership teams, and teachers indicated that the mock STW visits had the greatest impact on their improvement progress and goals.

Although originally intended as a learning opportunity for the projects schools, the mock STW visits also greatly benefited the overall project team (i.e., project leadership teams, state leadership). By visiting each school, the project team became more familiar with the challenges

project schools were facing, the early successes they were experiencing, and identified additional supports that would benefit the project schools. Ultimately, the mock STW visits impacted the project team's understanding of the project and informed future project-level discussions and decisions.

DISSEMINATION OF THE MLD PROJECT BEST PRACTICES

Even before STW designation, projects schools were being utilized as mentor schools hosting schools both from within their school district (including elementary and high schools), and schools outside of the district who wanted to visit and adopt the best practices that MLD Project schools had to offer. In addition, teachers and principals were presenting at state and national level conferences on their school's transformation due to their participation in the MLD Project.

> "People were asking all the time about our school! Directors of Instruction from other school districts wanted to come and see our school! Other schools in the district want to adopt some of the things we've put in place."
>
> - MLD Principal

Challenges

With any complex project, such as the MLD Project (i.e., 8 unique programmatic components, multiple states, multiple districts), a variety of challenges were bound to occur. Many of challenges faced by the MLD Project leadership team, state leaders, leadership coaches, and principals were monitored, discussed, and resolved. Other challenges, however, were more tenacious and strategies were implemented to ameliorate barriers to implementation when they could not be removed. Data collected for the project as well as minutes from project team (i.e., leadership team, state leaders, and evaluators) phone calls and meetings highlight the most challenging barriers.

EVOLUTION OF THE PRINCIPAL MENTORSHIP

The principal mentor was one programmatic components of the MLD Project that was never fully implemented as intended. Originally, each project school principal was to be paired with a principal mentor, who would meet in person with them 6 times a year. However, it was challenging to secure enough principal mentors for a one-to-one match for each of the 12 principals, and the frequency of 6 visits a year was prohibitive due to geographic distances between schools and time constraints. In the end, each state provided mentoring experiences for MLD Project principals through the following substitutions: principal mentors worked with principals virtually via phone or email in addition to in-person visits; principals received mentoring during visits to other STW schools or on STW designation visits; and principal mentors met with groups of principals at state-sponsored events. Given these substitutions, the project also reduced the amount of required mentoring to 4 sessions per year.

DISTRICT DEMANDS

Unfortunately, district demands often took precedence, and sometimes attention, away from the MLD Project focus. In particular, principals were sometimes laden with district changes, issues, and requirements that reduced their autonomy and decision making as leaders. For example, breaking up the middle-grades program across two buildings with the 6th grade housed in the elementary school and the 7th and 8th grade housed in the high school, which hindered the creation of a single, unified leadership team; tearing down the middle school building and relocating students and faculty into the high school, which created a loss of identity; sharing teachers between the middle school and the high school, which overburdened the teachers in 2 schools; removing a principal “effective immediately” with the superintendent assuming the role of interim principal; imposing districtwide responsibilities (i.e., Athletic Director, Assessment Director) in addition to the principal’s duties, which resulted in a conflict of interests; and needing to focus on meeting district demands, which led to the premature exit of the MLD Project for one

school. All of these challenges negatively impacted school culture. Leadership coaches employed responsive coaching to support the changing needs of schools during these transitions and to provide them with the resources that they needed.

In addition, the districts had their own professional development requirements for school attendance, which sometimes resulted in competing agendas, as well as made it difficult for principals to attend MLD Project professional development and networking opportunities. One of the ways that leadership coaches responded was to help schools to align these competing services, and manage the many changes, by making connections with the school's school improvement plan and the STW Rubric. This assisted many schools with staying focused on their vision for improvement and not becoming overwhelmed nor pulled in too many different directions.

Lessons Learned

There are a myriad of lessons learned from the MLD Project that are important to document as part of the evaluation of the implementation and outcomes of the project. These lessons provide valuable information for future implementation of this project and serve to inform other middle-grades schools embarking on leadership improvements and the implementation of a collaborative leadership model. Although there were many differences between the twelve project schools, such as their location (rural/small town/suburban), enrollment, building grade configuration, and district and state requirements, they shared many of the same challenges and struggles to improve their leadership practices, develop a student-centered focus, and improve outcomes. The lessons presented below reflect findings from the data collected as part of the evaluation from all stakeholders in the project including principals, leadership teams, teachers, coaches, principal mentors, and state project leaders.

VALUE OF PRINCIPALS WORKING WITH THEIR FACULTY TO ADOPT A GUIDING VISION

The STW criteria provided a vision to MLD Project schools for what they wanted their schools to become. Using the STW criteria and rubric helped schools to delve deeper through articulating the meaning of each criterion, how they could implement it successfully by identifying their school's strengths and weaknesses, and validating their existing best practices.

A guiding vision:

- Serves as a framework and a focus.
- Provides a picture and pathway of where the school is going.
- Brings the faculty together with a shared purpose.
- Serves as a basis for reflection and continuous improvement.
- Keeps the focus on student needs and student-centered learning.

> "I attribute the improvements we've made to establishing a vision. A lot of times you've got to be able to see the full picture of where you are going before you take the steps to get there."
> - MLD Principal

IMPACTFULNESS OF USING THE STW RUBRIC FOR CONTINUOUS IMPROVEMENT

Through the MLD Project, schools adopted and then facilitated their own continuous school improvement model where data was used at every stage of a never-ending cycle to: inform planning, set targeted goals, reflect on and evaluate progress, and refine implementation. The adoption of a continuous improvement structure began when schools were introduced to the STW

criteria as a vision and the STW Rubric as a tool to inform school decision making with a data-driven approach. This structured approach of using the STW Rubric in ongoing school improvement was critical to keep the focus of MLD Project schools on their goals and its existence will assist these schools in sustaining their work after the project ends.

Key aspects of how the STW Rubric facilitated continuous improvement:

- Best middle-grades practices based on evidence-based research.
- Endorses best practices that are already occurring in the school while creating a systemic approach to encourage consistency of practice and expectations.

> "The rubric helped to show us what is best for kids. And it was the evidence. It wasn't just this is what we feel is important. It is best practice, research based, so it helped in getting the buy-in."
>
> - MLD Principal

- Provides a common language for discussions about improvement leading to a shared vision.
- Provides a structure for setting classroom and building-level improvement goals
- Powerful tool for cultivating individual and school-wide reflective practices.
- Empowers teachers to accept the challenge of high performance and strive to continue to improve.
- Empowers teachers to focus on growing the "whole child" instead of a solitary emphasis on academic achievement.
- Tool to unify multiple programs the schools is required to implement with the school's shared vision to create one plan.
- Defines the school's identity as it was ingrained in the culture and informed every decision.

IMPORTANCE OF COLLABORATIVE LEADERSHIP

The MLD Project provided the structure and support to schools to facilitate the development of collaborative leadership. Project school principals now utilize a collaborative leadership style that empowers their teachers and provides them with opportunities for collaboration, reflection, and shared leadership through participation in the leadership team, interdisciplinary teams, PLCs, etc. More importantly, these collaborative leadership practices built school capacity to sustain the changes they were implementing. Time provided for collaboration, (e.g. PLCs, grade-level teams, interdisciplinary teams, etc.) often focused on the individual

student, such as monitoring early indicator data, goals, and challenges to identify ways to provide a support system for each student and to celebrate their successes. Project schools had begun "teaching smarter" as a result of their collaboration since they were sharing ideas, coordinating lessons, and looking more critically at their teaching practices.

Collaborative leadership empowers teachers to:

- Take on shared leadership roles.
- Increase their ownership for improvement goals and action plans.
- Build school capacity to sustain the changes over time, and sometimes despite leadership change.
- Provide internal professional development as a content expert

> "The principal has really empowered his staff, and that's when the school really started to take off. That's made a huge difference in momentum. He's still in charge, but he really has empowered different people in leadership roles and gotten many people involved. That has made a huge difference in that school in terms of cohesiveness and community."
>
> - MLD Leadership Coach

Key ingredients of collaborative leadership:

- Cultivate a collaborative culture between the administration and teachers.
- Develop the right person for the right job by encouraging teachers to take a lead in areas that interest them and by giving them support to try something new.
- Provide teacher leaders with decision-making power and autonomy.
- Develop structures for teachers to work together (e.g., leadership team, interdisciplinary teams, PLCs).
- Establish time for regular collaboration and reflection.
- Determine a common goal and action plan.
- Define roles for team members.

KEY FEATURES OF LEADERSHIP TEAMS

Leadership teams were already in place at several MLD Project schools before the project began, while at other school, they were established at the beginning of the project. All were organized as a result of implementing a collaborative leadership structure at each school. Although MLD Project schools used varying methods of structuring their leadership teams, organizing their activities, and establishing meeting frequency and duration, several key features of the teams were common across schools.

LESSONS LEARNED

Best practices for leadership team composition:

- Consider whether appointments or volunteers are best for your school.
- Goal is that members represent the whole faculty.
- Include a district representative.

Best Practices for leadership team organizational protocol:

- Establish group expectations and norms.
- Regular meeting schedule with agendas shared before meetings.
- Document decisions and next steps.

> "The best way to improve our climate was through the leadership team because they disseminated information to their departments about what we were planning and the positive approaches to addressing challenges."
>
> - MLD Principal

Strong leadership teams:

- Are engaged and focused on improvement.
- Examine data to drive actions.
- Have a continuous improvement approach.
- Communicate actively with the whole faculty.
- Build capacity for sustainability beyond the grant or principal turnover.
- Are constantly growing as they learn from challenges and setbacks.

TRANSFORMATIVE SUPPORT AND RESOURCES FOR PRINCIPALS

Through their participation in the MLD Project, principals received support and resources that they considered to be transformative to their leadership practices, beliefs, and behaviors. Principals reported that it was the combination of resources and supports that allowed them to work on the global culture and environment of their school, build collaborative leadership, work with teachers on instructional practices and interventions, and implement needed programs and practices (e.g., advisory, PLCs, discipline policy).

Key supports and resources for principals:

- Knowledge of middle-level philosophy, young adolescent needs, and best middle-grades practices.
- Networking opportunities with other principals.

- Self-assessment of leadership behaviors and skills and development of a professional growth plan
- Refining a leadership style built around collaborative practices.
- Visits to other high-performing schools.

> "I think for me the biggest piece was just to increase my knowledge about what works in the middle grades and how we can best implement those practices in my school setting."
>
> - MLD Principal

VALUE OF VISITS TO STW SCHOOLS

STW visits allowed groups of teachers from MLD Project schools to share a common experience by observing how a STW school had translated the STW Rubric's "best" practices into "everyday" practices. The visits were powerful as they allowed for a better understanding not only of the best practices described in the rubric's criteria, but the difference between implementing with "Good Quality" and "High Quality." The visits encouraged schools that they were already doing some best practices and once they began consistently implementing additional best practices and seeing positive results, their confidence increased.

Reasons why visits to STW schools were so impactful:

- Powerful to visit another school and see best practices in action.
- Validates existing best practices in project school.
- Gain knowledge about how another school is successful.
- Observe the difference between "good quality" versus "high quality."
- Faculty on visit shares a common experience.
- Take away the practices that could be adopted or modified to meet the project school's needs.

> "I feel that the visits to STW schools have provided a way for positive change to occur as the faculty sees other schools with similar populations and how they have made changes to bring about growth."
>
> - MLD Leadership Coach

- Encourages faculty to take ownership of improvement.
- Validates the practice of utilizing the STW Rubric as a vision and framework for the school
- Illustrates what the STW Rubric looks like in daily practice when implemented successfully

LESSONS LEARNED

IMPORTANCE OF NETWORKING

MLD Project schools valued the knowledge, resources, and support gained through networking opportunities and professional development provided at the state and national levels through the Forum's STW Network. For teachers, this meant increased opportunities to connect and learn from other teachers. Principals emphasized the benefits afforded to them from getting together with other principals to collaborate and share experiences and challenges. Project schools attended state and national middle-level conferences, as well as the annual national STW Conference. Rural schools especially valued these learning opportunities and viewed them as a lifeline from their rural isolation.

Value gained from networking with other schools:

- Opportunity to collaborate and share ideas.
- Creation of a professional development cohort between project schools.
- Facilitates a train the trainer model for professional development.
- Provides inspiration to make meaningful changes.
- Take away practices that could be adopted or modified for school.
- Gain knowledge about how another school is successful.
- Opportunities to present at state or national conference as part of the MLD Project.
- Builds school's expertise and confidence.

> "The grant provided the leadership team with supports to improve the capacity to instill change within the school and it provided us with opportunities to receive feedback. Additionally, the networking it provided has been invaluable and connected our school with other middle schools across the state. It has been the springboard to change for the benefit of our students and will create a lasting impact on the school."
>
> - MLD Principal

SCHOOL DISTRICT SUPPORT

The MLD Project built requirements into the structure of the project's activities and components that were intentionally designed to include district representatives in the process. The most visible requirement was that a

> "The superintendent bought in early on this process. He was at many of our meetings, when we were meeting with PLCs and with the leadership team. He would make time to come and be part of that, and he supported our principal in that school from day one and that's very critical. Everybody needs to get in on the game.
>
> - MLD Leadership Coach

district person be a regularly attending member of each school's leadership team, thus cultivating school district involvement and support of the project from the beginning and maintaining it throughout the project. Principals believed that through a better understanding of STW, the school district would recognize that the rubric captured all of the district's goals and vision.

School district support was impactful in meaningful ways:

- Importance of the school district providing the principal with the autonomy to make decisions and implement change.
- School district representation on the school leadership team.
- Cultivate continued school district support by:
    - Frequent communication or a presentation on the school's progress as a result of the grant to the school district;
    - Inviting superintendent participation in the school's mock visit with access to the expertise of the visiting national team of middle-grades experts; and
    - Inviting superintendent attendance to an AMLE presentation on the project or the STW Conference.

> "In the second year, my coach presented the progress that we had made and what to expect from the mock STW visit to the superintendent. Then the superintendent came to the mock STW visit and walked around with the visiting team. I think that really helped in seeing this as a free opportunity to make us better."
>
> - MLD Principal

- Emphasize school improvement supports would be provided at "no cost" to the school or school district.

BUILDING SCHOOL CULTURE

The first area of work that was addressed by the MLD Project leadership coaches was working with principals and leadership teams to build a positive, supportive, reflective, and student-centered school culture. The contextual changes in culture and climate were viewed as pre-requisites for leadership change, collaboration opportunities, and instructional improvements. In other words, school culture was the building block to improvements in other areas of the school.

Key ingredients when building school culture:

LESSONS LEARNED

- Transitioning from teacher-centered instruction to student-centered learning through student engagement and cooperative learning strategies.
- Creating consistent expectations and a better understanding of student behavior.
- Providing an adult advocate for each student and creating personal connections with students.
- Creating a supportive culture that allows risk-taking for teachers and students.
- Encouraging student investment in their own education.
- Developing a growth mindset among teachers and students.
- Empowering teacher leaders with decision-making power and autonomy.

> "We're moving from an adult-centered to a student-centered facility and programs. The piece that has really bought us growth from a teaching standpoint is active engagement strategies for youngsters in the classroom and that is really elevating the teaching and learning. That is critical. They absolutely engage kids. We've got a group of kids that have to be active, so might as well channel it in a useful way."
>
> - MLD Leadership Coach

- Scheduling time regularly for reflective conversations on teaching and building-level practices.
- Building a trusting environment to improve teaching practices through peer observations.
- Being intentional with school data to secure teacher buy-in for implementation.

BENEFITS OF SCHOOLS PARTICIPATING IN MOCK STW VISITS

MLD Project school principals, leadership teams, and teachers unequivocally indicated that their participation in mock STW visits at their school was the activity that taught them the most about their improvement progress and goals. Mock STW visits involved the MLD Project schools hosting a national team of STW-trained visitors whose purpose was to spend a day touring the school, meeting with administrators and teachers, observing classrooms, and interviewing students and parents. The trained teamed used the STW criteria as the framework for their observations and they provided the school with a comprehensive report of their strengths and challenges at the conclusion of the visit. Preparing for such a visit, participating in the visit, and receiving the feedback from the national team was positive and motivating experience.

Important elements of the mock STW visits to schools:

- Improved teaching practices as schools monitored teaching practices more critically utilizing the STW Rubric in preparation for the mock STW visit.

LESSONS LEARNED

- Received expertise from a national team of middle-grades professionals.
- Validated existing best teaching practices.
- Confirmed principal and leadership team assessments and decisions.
- Resulted in action plans for future improvements by reviewing feedback and addressing mediating questions to identify and prioritize growth areas.
- Helped schools to align their school improvement action plans with the STW Rubric.
- Utilized the mock STW visit process and feedback as a "pre-visit" to assess readiness for applying for a STW designation.
- Built school confidence that they could become a nationally recognized STW School.

BENEFITS OF PROJECT TEAM DOING MOCK VISITS FOR THEIR UNDERSTANDING OF SCHOOLS

The MLD Project team, including leaders from the National Forum, leaders from Kentucky and Michigan, and other project partners expressed the growth they experienced by visiting the MLD Project schools to conduct the mock STW visits. They described the experience as enlightening to see the schools firsthand and experience the challenges they face as well as the successful implementation of the MLD Project. This experienced and trained national team recommends this practice to other project teams because of its value to both the schools and the project leadership.

Important elements of the mock STW visits to project team:

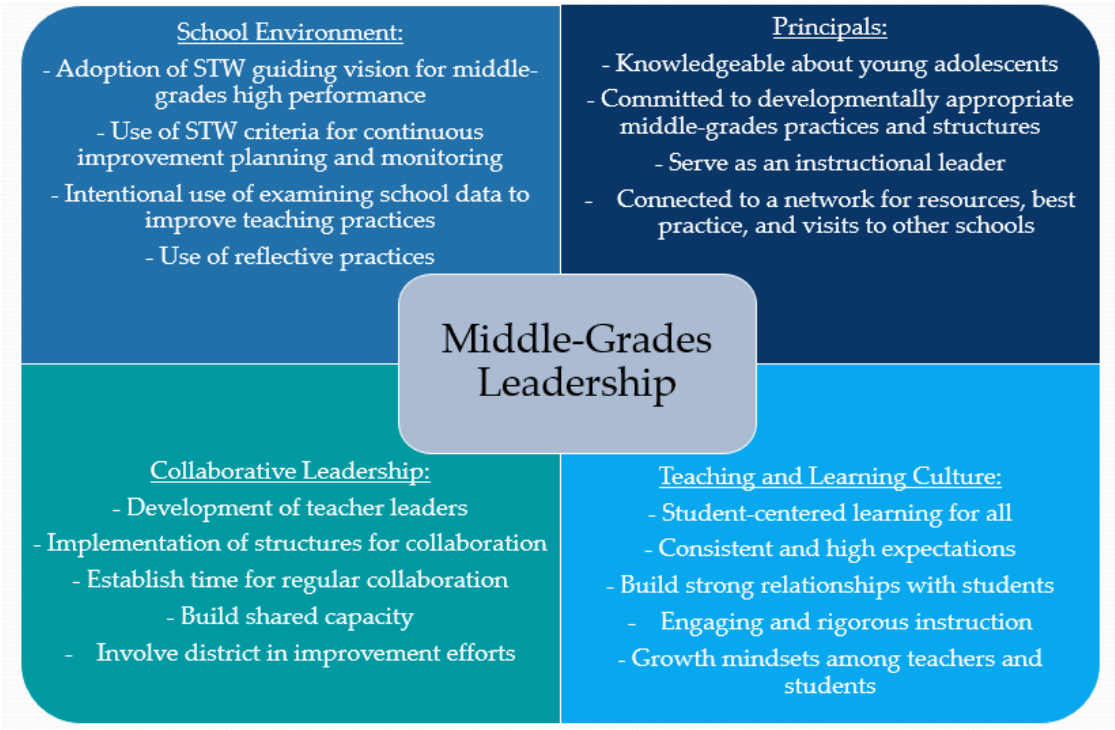
- Increased familiarity and ownership of all the schools participating in the MLD Project.
- Identification of similar challenges (and successes) across project schools for future project-level discussions and decisions.
- State directors developed a deeper understanding of their state's school through a fresh perspective from experienced colleagues.
- Increased familiarity of each state's expectations for their project schools and coaches.
- Built comradery between project team members through a shared experience.
- Resulted in a positive impact on the work of the project team.

KEY COMPONENTS OF PRINCIPAL EFFECTIVENESS

The results of the data analyses presented in this report are positive in that they highlight numerous improvements at MLD Project schools. The results also provide unique insight into a

middle-grades school improvement project focused on principal leaders and collaborative leadership. The figure below (Figure 7) depicts the key supports, activities, and practices implemented at MLD Project schools that were the most impactful on building middle-grades leadership effectiveness. They provide a road map for other middle-grades schools and principals that are struggling to improve. With key supports, activities, and practices at the school level, by the principal, through a collaborative leadership structure and in a positive teaching/learning culture, middle-grades leadership is more effective.

Figure 7. MLD Project Key Components of Middle-Grades Principal Effectiveness



Although the MLD Project results do not yet connect the contextual improvements in culture and climate, collaborative leadership practices, and the implementation of "best" middle-grades practices with student achievement growth, it is encouraging to find these improvements in programs and practices because improvements in these areas are the building blocks to improvements in student achievement. Student achievement cannot be improved without first changing and improving the school environment and the teaching and learning process. It is also encouraging to find student achievement growth at several MLD Project schools as this may indicate a trajectory toward more future growth. Finally, the fact that ten of the twelve project schools were designated as STW during the project or no-cost extension is another indicator of future achievement growth.

Sustainability

The project leadership team, state leaders, leaderships coaches, and principals involved in the i3 STW Project expressed a belief that project schools will not only sustain the improvements they have made, but continue to grow and improve, particularly in the area of student achievement. Sustainability by definition is the "capacity to endure" and although the evaluation cannot predict whether project schools will sustain their accomplishments and continue to grow, data collected on the project highlight several key structures and practices that are likely to support sustainability and continued growth.

STW DESIGNATION

As described previously and mentioned in the prior section, the framework for the MLD Project was the Forum's STW criteria. At a national level, the Forum uses the STW criteria for identifying middle-grades schools across the country that are on an upward trajectory of growth and continuous improvement. Schools that are interested in being designated as a STW by the Forum engage in an application process with the Forum's affiliate in their state. The application process is comprised of several steps, including: completing an online application form; submitting data on student practices and performance; completing the online STW rubric; and hosting a visit to their school by a STW review team that includes interviews with the administrators, teachers, students, and parents; observations of classrooms; and team meetings. Typically, each of the 18 STW state affiliates designates 3 to 5 schools per year as STW schools.

Although being designated as a STW was not articulated as an outcome for MLD Project schools by the end of the grant, it was understood that the project would lay the foundation for schools to be designated in the future. It turned out, however, that seven project schools made substantial improvements in practices and outcomes during the grant period and so applied and were designated as a STW school before the end of the grant. Three more project schools applied during the no-cost extension of the grant and were also designated as STW. They join more than 300 other schools across the country that are STW schools and they serve as success stories for the MLD Project. Because the STW process requires all schools to apply for re-designation every three years, sustainability is supported through a continuous improvement mindset. STW schools use the STW criteria and rubric to regularly examine and reflect on ways to continue to improve. As stated on the STW rubric "even when the ultimate goal is reached, a true high performing middle school will continue to seek ways to improve as new challenges arise." Sustaining the improvements and continuing to grow, therefore, is very likely among the MLD Project schools.

SUSTAINABILITY

The ten (of the twelve) project schools that were designated as Schools to Watch Schools by the National Forum to Accelerate Middle-Grades Reform are:

2017

- Garrard Middle School, KY
- Marshall Greene Middle School, MI
- Mayfield Middle School, KY
- Oscoda Middle School, MI
- Richfield Public School Academy, MI
- Rowan County Middle School, KY
- Union County Middle School, KY

2018:

- Grant Middle School, MI
- Mt. Morris Middle School, MI
- Reese Middle School, MI

> "I know whoever joins that team or leaves that team, that somebody's going to be on that team that says this is what we're doing that's great; it's working and we'd like to continue it. And I'm sure they're going to continue the work."
>
> - MLD Principal

SUSTAINABILITY PROTOCOL AND MATRIX

In thinking about how schools could sustain best practices and momentum beyond the grant, the MLD Project developed a sustainability protocol and matrix. The protocol included a three-step process: 1) discuss how to plan for sustainability; 2) complete the sustainability matrix; and 3) develop a plan for sustainability.

The first step was to begin a discussion about sustainability at the school utilizing an existing group (i.e., leadership team, interdisciplinary team, grade-level team, PLCs) or the entire faculty. This group would focus on why sustainability was important and what challenges might prevent sustainability in their school in order to strategize on what resources might be needed and how the school would evaluate success.

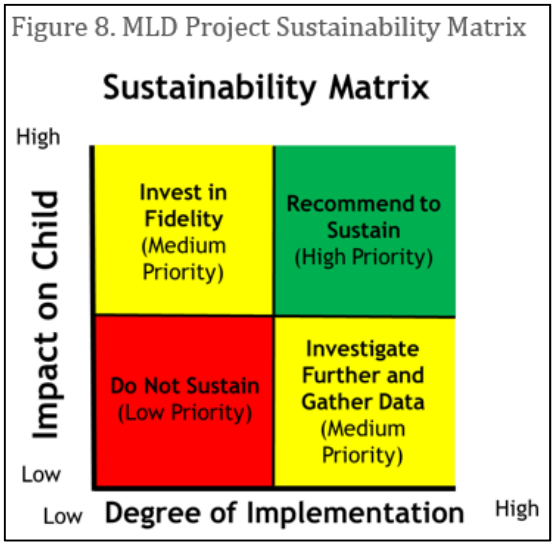
SUSTAINABILITY

> "Creating a sustainability matrix with the four domains of the STW Rubric to complete a plan for 2018 school year went very well. All faculty had a voice in determining next steps using the sustainability matrix as a tool. We were able to get a sense of what the faculty valued, what was not necessary to continue, and the next steps in what they need to continue to work on to advance their school forward."
>
> - MLD Leadership Coach

The second step was to complete the sustainability matrix. First, the group would identify all new practices they had implemented at their school as a result of their participation in the MLD Project. Then the group would identify any other practices and programs that were currently being implemented at their school, such as district initiatives, state mandates, or other voluntary initiatives to generate a comprehensive list. Finally, the group would discuss each practice (or program) on the list individually until they reached a consensus on the degree of implementation at the school and the level of impact on the child.

To facilitate this conversation, the group utilized the sustainability matrix as a graphic organizer with four quadrants based on the practice's impact on the child (high/low) and the degree of implementation at their school (high/low) to determine where to assign each practice (Figure 8). Each quadrant of the sustainability matrix was assigned a priority category of low, medium, or high with an accompanying recommendation. A high priority practice had a high impact on the child and was also high in the degree of implementation, and therefore the recommendation would be to sustain this practice. A practice that had a high impact on the child, but a low degree of implementation, received a medium priority and the recommendation to invest in the fidelity of implementation to improve the practice. A practice with a high level of implementation, but a low impact on the child, was also considered a medium priority, but the recommendation was to investigate further and gather data to determine if the practice should be sustained or discontinued. Finally, a practice that had both a low impact on the child and a low degree of implantation was considered a low priority with the recommendation to discontinue the practice. The intention of step two was to provide everyone in the group with a voice for a shared, reflective conversation about school practices, and the expectation was that some practices would be reassigned to a different quadrant after these discussions.

Figure 8. MLD Project Sustainability Matrix



SUSTAINABILITY

The third and final step was to develop a plan for sustainability by aligning the newly prioritized practices to the school's improvement plan. In this step the group revisited available resources and potential challenges to determine which practices they could sustain and then they incorporated those practices into their school improvement plan. Since continuous school improvement is the objective, targeted goals were to be set and monitored through data collection with plans to revisit the sustainability protocol with future school improvement planning.

The MLD Project leadership team provided training on how to conduct the sustainability protocol to project schools and leadership coaches. Then the leadership coaches worked with their individual schools to facilitate the protocol and develop a sustainability plan, which was submitted to the MLD Project leadership team before the end of the grant.

References

Flowers, N., & Mertens, S. B. (2003). Professional development for middle-grades teachers: Does one size fit all? In P. G. Andrews & V. Anfara, Jr. (Eds.), *Leaders for a movement: professional preparation and development of middle level teachers and administrators* (pp. 145-160). Greenwich, CT: Information Age Publishing.

Flowers, N., Mertens, S. B., & Mulhall, P. F. (1999). The impact of teaming: Five research-based outcomes of teaming. *Middle School Journal, 31*(2), 57-60.

Flowers, N., Mertens, S. B., & Mulhall, P. F. (2000a). What makes interdisciplinary teams effective? *Middle School Journal*, *31*(4), 53-56.

Flowers, N., Mertens, S. B., & Mulhall, P. F. (2000b). How teaming influences classroom practices. *Middle School Journal, 32*(2), 52-59.

Flowers, N., Mertens, S. B., & Mulhall, P. F. (2002). Four important lessons about teacher professional development. *Middle School Journal, 33*(5), 57-61.

Flowers, N., Mertens, S. B., & Mulhall, P. F. (2003). Lessons learned from more than a decade of middle grades research. *Middle School Journal, 35*(2), 55-59.

Flowers, N., Mertens, S.B., Mulhall, P.F. (2007). Applying current middle grades research to improve classrooms and schools. Westerville, OH: National Middle School Association.

Mertens, S. B., & Flowers, N. (2003). Middle school practices improve student achievement in high poverty schools. *Middle School Journal, 35*(1), 33-43.

Mertens, S. B., & Flowers, N. (2006). Middle Start's impact on comprehensive middle school reform. *Middle Grades Research Journal, 1*(1), 1-26.

Mertens, S. B., Flowers, N., Hesson-McInnis, M., & Bishop, C. (2006, April). *Middle grades teaching practices: A validation study*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Mertens, S. B., Flowers, N., Hesson-McInnis, M., & Bishop, C. (2007, April). *The impact of middle grade structures, organizations, and practices on student outcomes: Moving toward a multivariate/multi-level model*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Appendices

APPENDIX A. MLD PROJECT FIDELITY OF IMPLEMENTATION MATRIX

	Key Elements of Component	Operational Definition for Indicator	Implementation Levels
Component 1: Create a powerful vision for high performance using the STW criteria			
1	Orientation/meeting for schools	Orientation/meeting each year for school representatives and coaches (mentors optional) to: 1) Provide a project overview for the year; 2) Go over the STW criteria; 3) Introduce the support team; 4) Go over the responsibilities of each party; and 5) Engage in networking.	0 (low)=not attended 1 (med)=1 school representative attended 2 (high)=more than 1 school representative attended
2	STW Criteria training for schools	STW Criteria training each year for school representatives and district leaders from each school to provide an in-depth understanding of the STW Criteria and to continue to learn tools for using the criteria. Training may occur in a group setting or be conducted by coach at the school.	0 (low)=not attended 1 (med)=1 school representative attended 2 (high)=1 school representative and district representative attended
3	School revises/ establishes a vision	The MLD leadership coach works with the school leadership team to revise or establish their vision for high performance based on the STW criteria and a model of continuous improvement by the end of Year 1. In Years 2-3, the vision is revisited and revised as needed. Vision is written out and included in improvement plan.	0 (no)=not completed 4 (yes)=completed
4	Faculty supports the school vision	The degree to which the faculty supports the school vision via their response to a Self-Study Survey question (I support the vision/goals/ practices of the project).	0 (low)=<50% of faculty support vision 1 (med)=50-75% of faculty support vision 2 (high)=76%+ of faculty support vision
Implementation of component with fidelity= At least 70% of schools: Year 2 with med or high implementation (score of 1 or 2); Year 3 with high implementation (score of 2)			Range of 0-10 0 (low)=0-2 1 (med)=3-7 2 (high)=8-10

APPENDICES

	Key Elements of Component	Operational Definition for Indicator	Implementation Levels
Component 2: Engage in assessment and planning to identify needs, develop goals, and implement an action plan			
1	Faculty complete the STW Rubric	The STW Rubric is a tool for continuing improvement. It is divided into 4 sections depicting the STW criteria. Each section contains detailed examples of excellence. Teachers rate their implementation.	0 (low)=<50% faculty completed 1 (med)=50-75% faculty completed 2 (high)=76%+ faculty completed
2	Identify needs with school leadership team	The MLD leadership coach works with the leadership team to examine data (STW Rubric results, Self-Study Staff Survey results, student outcome data, etc.) and reflect on the STW criteria in order to identify areas for improvement.	0 (no)=not completed 1 (yes)=completed
3	Develop growth targets and action plan with school leadership team	Each year, the MLD leadership coach works with the leadership team to develop growth targets for each area of improvement, with an associated action plan for each. Schools follow state requirements for documentation to avoid duplicating efforts.	0 (no)=not completed 1 (yes)=completed
4	Faculty input and support of action plan	The leadership team presents the action plan to the faculty for input and support.	0 (no)=not completed 1 (yes)=completed
5	Implement the action plan	The MLD leadership coach works with the leadership team and faculty to implement the action plan. Performance benchmarks are included in the action plan in order to assess progress, monitor results, and make mid-course corrections. Approximately 2-4 goals per year.	0 (low)=<50% of action items completed 2 (med)=50-75% of action items completed 4 (high)=76%+ of action items completed
Implementation of component with fidelity= At least 70% of schools with high implementation (score of 2)			Range of 0-9 0 (low)=0-1 1 (med)=2-7 2 (high)=8-9

APPENDICES

	Key Elements of Component	Operational Definition for Indicator	Implementation Levels
Component 3: MLD leadership coach works with leadership team			
1	Orientation/training for MLD leadership coaches	Orientation/training meeting for MLD leadership coaches at the beginning of each year (or when a new coach is hired). Meeting agenda includes project overview, STW criteria overview, coaching activities and expectations, and calendar of project events.	0 (no)=not completed 1 (yes)=completed
2	MLD leadership coaches work with leadership team	Each year, MLD leadership coaches meet with the leadership team a minimum of once per month in person at the school from Sept to May.	*0 (low)=0-4 meetings per year* *2 (med)=5-7 meetings per year* *4 (high)=8 or more meetings per year*
3	MLD leadership coaches work with principal to develop a professional growth plan	Each year, MLD leadership coaches work with the principal to develop a professional growth plan for improving leadership skills based on Val Ed leadership skills assessment data.	0 (no)=not completed 4 (yes)=completed
4	MLD leadership coaches participate in ongoing training	Ongoing training includes participation in coach conference calls, networking via Edmodo or other mechanisms, state meetings, attendance at the national STW Conference, and going on STW visits.	0 (low)=0-3 events per year 1 (med)=4-5 events per year 2 (high)=6 or more events per year
Implementation of component with fidelity= At least 70% of schools with high implementation (score of 2)			Range of 0-11 0 (low)=0-2 1 (med)=3-9 2 (high)=10-11

	Key Elements of Component	Operational Definition for Indicator	Implementation Levels
Component 4: Principal mentor works with principal			
1	Principal mentors work with principals	Each state provides mentoring experiences for project principals. This can take the form of principal mentors who work with project principals by meeting with them (in person or via phone or email). It can also take the form of mentoring that occurs for project principals as part of visits to other schools, visits to STW, or going on STW designation visits. Mentoring is expected to occur at least four times per year.	*0 (low)=0-1 meetings per year* *2 (med)=2-3 meetings per year* *4 (high)=4 or more meetings per year*
Implementation of component with fidelity= At least 70% of schools with high implementation (score of 2)			Range of 0-4 0 (low)=0 1 (med)=2 2 (high)=4

	Key Elements of Component	Operational Definition for Indicator	Implementation Levels
Component 5: Network school works with school			
1	School visits network school	Network schools host visits from the principal and leadership team members, providing observation opportunities (e.g., grade level meetings, team meetings, classrooms), discussion opportunities, and share resources. This manifests with either visits to STW sites, visits to or attendance on STW designation visits, visits to other i3 schools, or visits to other high performing schools. Schools go on a minimum of 2 visits per year.	*0 (low)=no visits* *2 (med)=1 visit* *4 (high)=2 or more visits*
Implementation of component with fidelity= At least 70% of schools with med or high implementation (score of 1 or 2)			Range of 0-4 0 (low)=0 1 (med)=2 2 (high)=4

	Key Elements of Component	Operational Definition for Indicator	Implementation Levels
Component 6: School engages in a leadership professional learning community			
1	School establishes a leadership team	Each school establishes a leadership team composed of a minimum of 6 people including the principal, district administrator, 3 teacher leaders, and one other staff member (e.g., counselor, assistant principal, curriculum coordinator).	*0 (low)=not established* *1 (med)=established, but not with 6 members per definition* *2 (high)=established with all 6 members per definition*
2	Training for leadership team on functioning as a professional learning community (PLC)	Training at the beginning of each year for leadership team members. Training agenda includes definition of a PLC (collegial group of administrators and teachers who work together regularly to collaborate and provide leadership for the school), operational procedures, group norms, and communication mechanisms.	*0 (no)=not completed* *1 (yes)=completed*
3	Regular attendance of leadership team members at leadership team meetings	Leadership team members attend meetings regularly (twice per month in Oct, Jan, Feb, Mar, Apr, and May; once per month in Sep, Nov, and Dec.), with few absences.	*0 (low)=Average attendance at meetings is less than 50% of members* *2 (med)=Average attendance at meetings is 50-75% of members* *4 (high)=Average attendance at meetings is 76-100% of members*
Implementation of component with fidelity= At least 80% of schools with high implementation (score of 2)			Range of 0-7 0 (low)=0-2 1 (med)=3-5 2 (high)=6-7

APPENDICES

	Key Elements of Component	Operational Definition for Indicator	Implementation Levels	
Component 7: School participates in networking experiences				
1	School attends national STW Conference	At least one person from the school attends the annual national STW Conference, attending sessions based on the areas for improvement identified as part of developing the action plan. Information is brought back to the school and reported to the full leadership team.	*0 (low)=not attended* *1 (med)=attended, but information not shared with leadership team* *2 (high)=attended and information shared with leadership team*	
2	School attends i3 MLD meeting at national STW Conference	At least one person from the school attends the i3 MLD project meeting at the annual national STW Conference. The meeting, hosted by the state project directors, is designed to develop collaborative relationships between schools and promote discussion of project challenges and successes.	Year 2: 0 (no)=not attended 1 (yes)=attended Year 3: NA	
3	School attends the state middle school conference or regional middle school workshop	At least one person from the school attends the state middle school conference (Michigan) or regional middle school workshop (Kentucky)	0 (no)=not attended 1 (yes)=attended	
Implementation of component with fidelity= At least 70% of schools with high implementation (score of 2)			Year 2: Range of 0-4 0 (low)=0 1 (med)=1-2 2 (high)=3-4	Year 3: Range of 0-3 0 (low)=0 1 (med)=1-2 2 (high)=3

	Key Elements of Component	Operational Definition for Indicator	Implementation Levels
Component 8: School receives regular and systematic professional development trainings based upon needs and data from the evaluation process			
1	School leadership team attends a professional development training on leadership, collaboration, and STW	Each year, state project teams or coaches conduct a professional development session focused on leadership, collaboration, and STW with a focus on continuous improvement. At least 2 people from the school's leadership team attend the session.	0 (no)=not attended 1 (yes)=attended
2	Schools participate in professional development focused on their identified needs	MLD leadership coaches and state teams provide/arrange for professional development focused on the school's identified needs, such as student engagement, differentiated instruction, technology, and creating a personalized learning environment. A minimum of 2 sessions per year with at least 3 to 5 faculty from the school attends.	*0 (low)=0 sessions attended* *1 (med)=1 session attended* *2 (high)=2 or more sessions attended*
Implementation of component with fidelity= At least 70% of schools with high implementation (score of 2)			Range of 0-3 0 (low)=0-1 1 (med)=2 2 (high)=3

APPENDIX B. MLD PROJECT K-PREP AND M-STEP ACHIEVEMENT FOR OVERALL STATE AND PROJECT SCHOOLS BY YEAR

Reading K-PREP Results 2015-2017	% of Proficient or Advanced			Growth 2015 to 2017
School	2015	2016	2017	
Overall State	53.8%	55.2%	56.9%	3.1%
School 1	39.2%	43.9%	49.3%	10.1%
School 2	53.5%	48.1%	48.6%	-4.9%
School 3	44.1%	44.4%	48.3%	4.2%
School 4	50.9%	49.5%	47.8%	-3.1%
School 5	58.0%	57.5%	58.7%	0.7%
School 6	46.1%	45.2%	45.3%	-0.8%
6th Grade				
Overall State	52.9%	55.5%	58.9%	6.0%
School 1	30.0%	43.2%	54.7%	24.7%
School 2	57.5%	47.2%	46.4%	-11.1%
School 3	45.0%	42.4%	48.2%	3.2%
School 4	44.6%	50.8%	48.5%	3.9%
School 5	56.1%	60.1%	52.7%	-3.4%
School 6	45.3%	38.0%	56.9%	11.6%
7th Grade				
Overall State	54.5%	56.6%	54.6%	0.1%
School 1	44.9%	42.3%	48.4%	3.5%
School 2	47.2%	52.7%	48.3%	1.1%
School 3	43.4%	48.1%	45.6%	2.2%
School 4	52.2%	49.5%	47.4%	-4.8%
School 5	55.2%	54.1%	57.7%	2.5%

School 6	46.5%	46.6%	35.0%	-11.5%
8th Grade				
Overall State	54.1%	53.6%	57.1%	3.0%
School 1	42.5%	46.0%	45.3%	2.8%
School 2	55.9%	44.0%	50.9%	-5.0%
School 3	43.9%	43.4%	51.5%	7.6%
School 4	57.6%	48.0%	47.3%	-10.3%
School 5	62.9%	58.4%	66.5%	3.6%
School 6	46.3%	50.9%	47.9%	1.6%
	Higher than the overall state			

Mathematics K-PREP Results 2015-2017	% of Proficient or Advanced			Growth 2015 to 2017
School	2015	2016	2017	
Overall State	42.8%	47.0%	47.0%	4.2%
School 1	36.7%	40.8%	44.7%	8.0%
School 2	34.8%	40.9%	31.1%	-3.7%
School 3	24.7%	31.3%	28.4%	3.7%
School 4	37.1%	37.2%	38.1%	1.0%
School 5	51.6%	58.4%	58.8%	7.2%
School 6	38.3%	39.7%	36.3%	-2.0%
6th Grade				
Overall State	43.2%	50.2%	49.1%	5.9%
School 1	25.0%	42.7%	52.2%	27.2%
School 2	37.4%	40.0%	27.6%	-9.8%
School 3	25.2%	35.4%	32.9%	7.7%
School 4	38.0%	43.1%	41.2%	3.2%
School 5	54.4%	70.0%	56.0%	1.6%

School 6	39.5%	30.4%	40.4%	0.9%
7th Grade				
Overall State	40.9%	45.4%	43.3%	2.4%
School 1	38.0%	36.6%	42.0%	4.0%
School 2	29.6%	38.6%	31.5%	1.9%
School 3	28.9%	35.3%	24.1%	-4.8%
School 4	38.9%	36.9%	36.8%	-2.1%
School 5	49.3%	55.9%	61.0%	11.7%
School 6	38.8%	40.2%	31.3%	-7.5%
8th Grade				
Overall State	44.2%	45.5%	48.7%	4.5%
School 1	46.8%	42.9%	40.7%	-6.1%
School 2	37.4%	44.0%	34.0%	-3.4%
School 3	19.7%	24.1%	27.6%	7.9%
School 4	33.7%	31.5%	35.7%	2.0%
School 5	50.9%	49.8%	59.9%	9.0%
School 6	36.8%	48.5%	38.7%	1.9%
	Higher than the overall state			

ELA M-STEP Results 2015-2017	% of Proficient or Advanced			Growth 2015 to 2017
6th Grade	2015	2016	2017	
Overall State	44.7	45.0	43.6	-1.1%
School 7	37.8	45.9	45.7	7.9%
School 8	56.8	56.1	45.2	-11.6%
School 9	20.0	20.0	22.2	2.2%
School 10	42.6	28.9	40.0	-2.6%

School 11	41.4	35.2	23.2	-18.2%
School 12	23.0	24.6	14.7	-8.3%
7th Grade				
Overall State	49.1	47.1	44.8	-4.3%
School 7	45.1	46.8	50.0	4.9%
School 8	59.0	51.6	54.0	-5.0%
School 9	43.9	23.8	22.3	-21.6%
School 10	50.0	37.8	24.7	-25.3%
School 11	48.1	41.0	29.7	-18.4%
School 12	27.9	22.4	19.7	-8.2%
8th Grade				
Overall State	47.6	48.8	48.0	0.4%
School 7	50.4	46.7	51.4	1.0%
School 8	51.9	53.2	58.4	6.5%
School 9	33.8	31.7	28.6	-5.2%
School 10	41.5	57.1	47.4	5.9%
School 11	43.9	46.3	50.0	6.1%
School 12	30.8	25.4	32.2	1.4%

Higher than the overall state

Mathematics M-STEP Results 2015-2017	% of Proficient or Advanced			Growth 2015 to 2017
6th Grade	2015	2016	2017	
Overall State	33.3	32.8	34.2	0.9%
School 7	23.9	32.6	27.5	3.6%
School 8	27.2	30.6	37.3	10.1%
School 9	12.1	6.5	6.5	-5.6%

School 10	24.5	20.5	30.3	5.8%
School 11	29.3	29.4	19.6	-9.7%
School 12	9.5	11.8	<5	-4.5%
7th Grade				
Overall State	33.3	35.3	36.2	2.9%
School 7	36.9	30.7	35.7	-1.2%
School 8	37.6	43.7	44.1	6.5%
School 9	17.7	10.0	8.5	-9.2%
School 10	30.4	29.9	26.0	-4.4%
School 11	34.0	29.5	39.1	5.1%
School 12	10.1	10.3	<5	-5.1%
8th Grade				
Overall State	32.2	32.7	33.5	1.3%
School 7	37.0	26.2	29.9	-7.1%
School 8	34.6	40.6	50.8	16.2%
School 9	6.8	8.4	7.7	0.9%
School 10	35.4	39.4	35.1	-0.3%
School 11	38.6	29.6	28.1	-10.5%
School 12	23.1	9.9	13.6	-9.5%

Higher than the overall state

APPENDIX C. MLD PROJECT IMPACT STUDY MODEL FOR ELA/READING

```
Computing standard errors:

Mixed-effects REML regression                   Number of obs     =      5,712
Group variable: Sch_ID                          Number of groups  =         50

                                                Obs per group:
                                                              min =         31
                                                              avg =      114.2
                                                              max =        190

                                                Wald chi2(15)     =    6303.57
Log restricted-likelihood = -5668.0208          Prob > chi2       =     0.0000
```

ReadSS_Z8	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Intervention	.043147	.0551131	0.78	0.434	-.0648727	.1511667
ReadSS_Z5	.6137239	.0096427	63.65	0.000	.5948245	.6326233
AA	-.2508199	.0408064	-6.15	0.000	-.3307989	-.1708408
Hisp	-.0305184	.0384307	-0.79	0.427	-.1058413	.0448045
Other	.0342738	.0401967	0.85	0.394	-.0445102	.1130578
female	.2182965	.0172052	12.69	0.000	.184575	.252018
MI	-.2103298	.0976617	-2.15	0.031	-.4017433	-.0189163
LEPStatus	-.0277801	.0745085	-0.37	0.709	-.1738141	.1182539
Special_Ed	-.3411173	.0302466	-11.28	0.000	-.4003995	-.2818351
Economically_Disadvantaged	-.2016305	.0187118	-10.78	0.000	-.2383049	-.1649561
Sch_AMO_AYP_c	.0203617	.0063698	3.20	0.001	.0078771	.0328463
Sch_Enroll_c	.0002935	.0001826	1.61	0.108	-.0000645	.0006514
Sch_FRL_c	-.002464	.0026868	-0.92	0.359	-.00773	.0028021
Sch_AA_c	.0070514	.0032808	2.15	0.032	.0006211	.0134816
Sch_Hisp_c	.0004671	.0042532	0.11	0.913	-.007869	.0088032
_cons	.1224304	.048695	2.51	0.012	.02699	.2178709

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
Sch_ID: Identity				
var(_cons)	.0213914	.0056025	.0128028	.0357414
var(Residual)	.4122673	.007754	.3973465	.4277484

LR test vs. linear model: chibar2(01) = 152.77 Prob >= chibar2 = 0.0000

APPENDIX D. MLD PROJECT IMPACT STUDY MODEL FOR MATH

```
Computing standard errors:

Mixed-effects REML regression                   Number of obs      =      5,737
Group variable: Sch_ID                          Number of groups   =         50

                                                Obs per group:
                                                              min =         31
                                                              avg =      114.7
                                                              max =        196

                                                Wald chi2(15)      =    7278.98
Log restricted-likelihood = -5190.5025          Prob > chi2        =     0.0000
```

MathSS_Z8	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Intervention	-.0004777	.0816594	-0.01	0.995	-.1605271	.1595717
MathSS_Z5	.6792261	.0092622	73.33	0.000	.6610726	.6973797
AA	-.0934549	.0372165	-2.51	0.012	-.1663978	-.0205119
Hisp	-.0525098	.0348092	-1.51	0.131	-.1207345	.0157149
Other	.0324889	.0366206	0.89	0.375	-.0392863	.104264
female	.1430595	.0156584	9.14	0.000	.1123697	.1737493
MI	-.1913631	.1447356	-1.32	0.186	-.4750397	.0923135
LEPStatus	.100634	.0667853	1.51	0.132	-.0302629	.2315308
Special_Ed	-.1559898	.0272821	-5.72	0.000	-.2094618	-.1025179
Economically_Disadvantaged	-.1509614	.0170786	-8.84	0.000	-.1844348	-.1174881
Sch_AMO_AYP_c	.0235668	.0094069	2.51	0.012	.0051296	.042004
Sch_Enroll_c	.000293	.0002702	1.08	0.278	-.0002367	.0008226
Sch_FRL_c	-.0027847	.0039545	-0.70	0.481	-.0105354	.004966
Sch_AA_c	-.002999	.004808	-0.62	0.533	-.0124224	.0064245
Sch_Hisp_c	-.0018937	.0062427	-0.30	0.762	-.0141292	.0103418
_cons	.1042299	.0708356	1.47	0.141	-.0346053	.2430651

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
Sch_ID: Identity				
var(_cons)	.0526652	.0122423	.0333932	.0830595
var(Residual)	.3434574	.0064453	.3310543	.3563252

```
LR test vs. linear model: chibar2(01) = 543.00        Prob >= chibar2 = 0.0000
```