International Journal of Teacher Education and Professional Development
Volume 7 • Issue 1

Middle School Mathematics Achievement:
Effects of Math Teacher Professional Development

Gisele Ragusa, University of Southern California, USA*
https://orcid.org/0000-0003-3056-5494

Shaobo Huang, University of Saskatchewan, Canada
https://orcid.org/0000-0003-1860-9925

Svetlana Levonisova, Chapman University, USA

ABSTRACT

Teacher professional development is often proposed as a means for improving students' achievement; however, few studies have been successful in empirically connecting teacher inservice interventions to their students' achievement gains, and especially in mathematics. The research presents results of a longitudinal study of an inservice teacher professional development that had as its primary purpose to improve middle school students' mathematics achievement. The study utilized a cross-school comparative research approach for the purpose of examining students' math achievement trajectories Hierarchical linear modeling was used for the study's data analyses. Results of the study revealed that the teacher professional development intervention had positive impacts on both the participating teachers and their middle school students. The participating students' achievement improved significantly for those whose teachers participated in the teacher professional development intervention.

KEYWORDS

Mathematics Achievement, Mathematics Teaching, Middle School, Teacher Professional Development

BACKGROUND, CONTEXT, AND PURPOSE

Teacher professional development is often proposed as a means for improving students' achievement; however few studies have been successful in empirically connecting teacher inservice interventions to their students' achievement gains, especially in mathematics. This paper presents results of a longitudinal study of a three-year inservice teacher professional development program that had as its primary purposes to improve middle school students' mathematics achievement and motivation for engaging in math. The study utilized a cross-school comparative research approach to examine students' math achievement trajectories, employing a quasi-control condition in the first year of the program's implementation and multiple rounds of inservice teacher professional development

DOI: 10.4018/IJTEPD.337965 *Corresponding Author

interventions in the remaining years. Hierarchical linear modeling was used for the study's data analyses with multiple iterations of model fitting.

RESULTS AND FINDINGS

Results of the three-year study revealed that the teacher professional development intervention had positive impacts on both the participating teachers and their middle school students. The teachers' efficacy for teaching mathematics increased for all participating teachers. Both students' intrinsic and extrinsic motivation were found to positively affect their mathematics achievement, which was measured by standardized math achievement tests and the participating district's mathematics benchmark examinations. The participating students' achievement improved significantly for those whose teachers participated in the teacher professional development intervention.

BRIEF SUMMARY, CONCLUSIONS, AND IMPLICATIONS

In summary, the participating students' achievement in this study improved overall, with greater increases resulting from increased exposure to the teacher intervention across time. This research has positive implications for middle to large size secondary schools in terms of the impact of teacher professional development on students' achievement as it had multiple positive effects on both the participating teachers and their students.

MIDDLE SCHOOL MATHEMATICS ACHIEVEMENT: EFFECTS OF TEACHER PROFESSIONAL DEVELOPMENT

The United States has become increasingly concerned about future workforces, so much so that from 2009 onward, science, technology, engineering, and mathematics (STEM) education has been a major focus of the U.S. federal administration's plans for secondary education. STEM education in K-12 involves the inclusion of technology and engineering in math and science academic programs. Specifically, technology develops skills and abilities in adaptability, complex communication, non-routine problem solving, and systems thinking to "shape our material, intellectual, and cultural world" (Bybee, 2010, p. 31; see also Morgan & Morgan, 2013). These skills and abilities are of primary concern to the team responsible for this research endeavor.

The utilization of curriculum as one of the primary components that prepare students for future workforces in STEM is undoubtedly an effective approach to introduce students to the skills and abilities needed in STEM professions. Various pedagogical approaches have been integrated in curriculum design to improve students' mathematics skills and abilities, such as problem-based learning and reality-focused instruction (Hansen & Gonzalez, 2014; Uygun & Tertemiz, 2014; Fulton, 2012). To design reality-based curriculum, or that which is focused on realistic scenarios embedded in students' lived experiences, contemporary disciplinary content knowledge in addition to pedagogical knowledge and teaching efficacy are necessary (Hashweh, 2009). However, reviews of student achievement data in middle school mathematics call into the question math teachers' disciplinary content knowledge, pedagogical knowledge, and teaching efficacy (Phelps & Howell, 2016; Thanheiser et al., 2010).

This study is part of a broader three-year research project that investigated the effects of a comprehensive teacher professional development intervention on middle school students' mathematics achievement. The enabling structures included in this intervention were: (a) University-based research laboratories; (b) a teacher training leadership team; (c) a content expert math advisory team; (d) math teacher professional development using a summer teacher academy and associated follow-up; (e) use of teacher-centric lesson study; (f) use of diagnostic teaching, inquiry focused math, and data driven decision making; and (g) a focus on integrated mathematics. Specifically, the purpose of this study was to determine the impact of teacher professional development on teachers' efficacy in teaching mathematics, their understanding of math content, and, ultimately, on their students' math achievement. Accordingly, the research explores the relationship between teacher and student factors and student achievement using multilevel, hierarchical linear modeling.

In this research, the participating teachers received either two or three years of professional development (the treatment), and therefore their sixth- through eighth-grade students received either a single, double, or triple year "dose" of the teacher intervention depending on the year that they entered middle school. The specific teacher interventions within the professional development program consisted of: use of academic language in mathematics contexts, instruction to improve content area literacy and use of informational texts, applied mathematics and real-life problem solving, strategies for effective and efficient use of informational math texts, strategies for improvement of content area literacy, effective uses of technology in the classroom, a teacher "boot camp" approach to improving mathematics content knowledge, and strategies for nimble data driven lesson design focused on inquiry and the learning cycles. The teachers had annual weeklong summer teaching academies followed by grade level and subject specific mini camps during their academic year at school sites and in grade level teacher study groups. Lesson study was a major component of the teachers' professional development. A typical yearlong dose of professional development included 30 hours of teacher summer academy, followed by 15-20 hours per semester during the academic year in lesson study workshop groups (50-60 professional development hours per year =1 dose).

REVIEW OF THE LITERATURE

The review of research that follows not only addresses the significant challenges and needs of mathematics teachers and their students, it also fully informed the design, development and impact testing of the teacher professional development intervention presented herein. Accordingly, the intervention addressed teacher content needs via summer teaching academies with university level STEM content experts, included content-to-pedagogical disciplinary mentored practice, involved teacher self-study of their lessons (Kayapinar, 2016; Stigler & Hiebert, 2009), and addressed students' math achievement and motivation for mathematics.

Teacher Efficacy and Its Impact on Teacher Performance

Teacher efficacy reflects a teacher's perceptions of their ability to "bring about desired outcomes of student engagement and learning, even among those students who may be difficult or unmotivated" (Tschannen-Moran & Hoy, 2001, p. 26). It can be derived from performance accomplishments, vicarious experiences, verbal persuasion, and emotional arousal and consists of three key components: efficacy for instructional strategies, efficacy for classroom management, and efficacy for student engagement (Lazarides, 2021; Klassen et al., 2011). Teachers' beliefs about their self-efficacy have been identified as critical factor in teachers' effectiveness (Nurindah et al., 2019; Riggs & Enochs, 1990).

Previous research has documented a positive link between teaching efficacy and teacher performance in the classroom and beyond. For example, Tschannen-Moran & Hoy (2001) measured teacher efficacy and its correlation with teacher professional behaviors. Results of this research indicated significant correlations between teaching efficacy and instructional strategies, classroom management, and student engagement. Other studies found a relationship between teacher efficacy and teacher adoption of innovative teaching (Fuchs et al., 1992; Ghai & Yaghi, 1997). Notably, investigating the relationships among teachers' experience, efficacy, and attitudes toward the implementation of instructional innovation, Ghai and Yaghi (1997) found that experience was negatively correlated and personal teaching efficacy was positively correlated with teachers' attitudes toward implementing new instructional practices. Moreover, empirical findings have revealed that teachers' efficacy plays a key role in influencing and sustaining teachers' professional commitment to teaching (Coladarci, 1992) and their job satisfaction (Caprara et al., 2006). Teacher efficacy was also found to be significantly related to various teacher beliefs, such as perceived academic climate, and teacher wellness (Klassen et al., 2008; Protheroe, 2008). Finally, it has been determined that teachers with a low sense of efficacy were most likely to drop out of the teaching profession (Glickman & Tamashiro, 1982).

Furthermore, not only does teacher efficacy affect teacher-related outcomes, it also impacts students' achievement, and positive teaching efficacy can lead to student gains in the classroom (Moore & Esselman, 1992; Ross, 1998). Tschannen-Moran and Barr (2010) found significant positive relationships between collective teacher efficacy and student achievement on standardized eighth-grade mathematics tests. Furthermore, teacher efficacy has been associated with enhanced student motivation (Roeser et al., 1993; Mojavezi, & Tamiz, 2012) and increased student self-esteem (Borton, 1991).

In a study conducted by Brown et al. (2021), 81 pre-service teachers who were preparing to teach children in preschool through grade 6 were studied to determine the extent to which student teaching impacted pre-service teachers' perceptions of preparedness to teach and sense of teaching efficacy and to explore the relationship between pre-service teachers' perceptions of preparedness to teach, their sense of teaching efficacy, and their ability to perform teaching tasks during student teaching. A retrospective casual-comparative research design was used for this study with several instruments: measuring perception of pre-service teachers' preparedness (MPPTP), teacher sense of efficacy scale (TSES), and a benchmark survey. A paired sample t-test, two-tailed Pearson R, and multiple regression were used to analyze the relative impact of scores on the MPPTP and TSES for pre- and post-student teaching. This study's results revealed a statistically significant mean difference ($p<0.001$) between pre- and post-assessment on teachers' pedagogical content knowledge, planning and preparation, classroom management, promotion of family involvement, and professionalism. The study also showed a statistically significant positive correlation between the preservice teachers' feelings of preparedness to teach and their ability and teaching performance ($r=0.35$). Hands-on student teaching experiences were found to increase pre-service teachers' sense of teaching efficacy and feeling of preparedness to teach.

Accordingly, teachers' beliefs about their teaching play a unique role in setting standards and cultivating the conditions which ultimately foster their students' academic success. Teachers who believe they can teach all students in a way that enables them to obtain optimal scholastic attainments, are more likely to exhibit teaching behaviors that support this goal and contribute to their students' academic success (Protheroe, 2008). Therefore, the present study's intervention was designed to provide efficacy-building experiences for math teachers through professional development opportunities and associated activities.

Needs for Intensity and Multi-Dimensional Structures in Teacher Professional Development

Teacher professional development research has also indicated that particular intensity and structures are best suited for mathematics professional development. Importantly, inservice teacher professional development needs to be sustained and intensive for it to translate into student achievement gains (Wei et al., 2009; Darling-Hammond, 2012; Hill, & Ball, 2004). Through a meta-analysis of professional development interventions, Wei and colleagues (2009) found that programs offering 30-100 teacher professional development contact hours over six to 12 months had a significant effect on student achievement gains. For programs offering an average of 49 hours in one calendar year, student achievement rose by 21 percentile points (Wei et al., 2009). Similar results were found in Hill and Ball's research (2004). Hill and Ball's hierarchical linear model (HLM) developed using data from 398 teachers in 15 institutes indicated that the longer the professional development, the more teachers learned and the higher the gains students achieved in mathematics.

In a study conducted by Hanuscin and colleagues (2021), a cross-disciplinary and collaborative, long-term shadowing model was proposed as a form of professional development to impact novice and new teachers in learning across cognitive, practical, relational, and emotional domains. Nine faculty members from Western Washington University experienced a shadowing model that originated from the North Cascades Olympic Science Partnership. Through self-study and narrative inquiry, the authors used vignettes from teachers' stories to highlight how shadowing helps support the professional development of teachers. Unlike an evaluation observation, shadowing provides contextually relevant and personalized learning opportunities for both the shadower and those who are shadowed. The flexibility in the shadowing model structure allowed teachers to shadow different teachers across disciplines and extend their experiences with debriefing, reflection, discussion, and co-facilitation. Some teacher participants believed that "the shadowing experience provided opportunities to build a mentor-mentee relationship," which prolonged long-term engagement in professional learning compared to traditional teacher professional development (Hanuscin et al., 2021; 38). The shadowing experiences also allowed teachers to focus on cognitive development, where they can contextualize their knowledge and develop new pedagogies from observation. The teachers in Hanuscin and colleagues' research were able to build upon their rational and emotional development to help their self-understanding and awareness of their educator identity. This prolonged shadowing experience has promise as a teacher professional development component across contexts in educational settings.

In research by Garet et al. (2010) it was determined that teachers' instructional practices can also be effectively improved by particular structures and content of professional development. In their study of the impact of professional development on teacher knowledge, teacher instructional practices, and student achievement in math, 84 seventh-grade teachers received professional development through summer institutes and seminars. The curriculum provided by in this research focused on chapter organization, lesson components, and supported instructional approaches, and it had an emphasis on subject-specific content. The described professional development program produced a statistically significant high impact on the participating teachers' instructional practices. Moreover, the frequency with which the teachers engaged in activities that elicited student thinking was shown to be higher for the teachers who engaged in the teacher intervention than for those who did not. Garet and colleagues (2008) also identified that teacher institutes with research-based content and follow-up seminars led to improved teacher practice. Moreover, teachers who attended both institute series and school coaching used more effective pedagogical practices than their peers. These results indicate that sustained and intensive professional development with collaborative structures can have a significant impact on student academic outcomes. It is for this reason that the described teacher intervention included both a summer immersion component of 60 + hours and ongoing fall and spring semester collaborative follow-up.

Lesson Study and Its Link to Teaching Effectiveness

The effectiveness of teacher professional development programs varies with certain types of structures, and some have been found to yield more impactful results. For example, lesson study, according to Stigler and Hiebert (2009), refers to a professional development process whereby teachers closely self-examine their lessons with a focus on addressing student needs via data-driven decision making, creating powerful and relevant curricula, and redesigning lessons. This approach has been successful in improving teaching effectiveness. Lesson study, per Stigler and Hiebert (20069, consists of three phases: collaboration, co-planning, and observing actual lessons with a focus on student thinking. In the lesson study model, teachers learn together. They plan, observe, and refine lessons to achieve their long-term goals for student learning and achievement. An essential component of lesson study is observing and teaching lessons, which are improved collaboratively. This compels teachers to examine their practice in depth in the context of student learning, connects them with their students and their professional community, and inspires them to engage in continuous instructional improvement. This model of teacher professional development has been applied widely and successfully in Japan, where it is recognized as a system for generating professional knowledge about teaching (Rozimella, 2020; Hiebert et al., 2002). Lesson study has recently been initiated by teachers at many sites across the

United States and is especially applicable to mathematics education (Ragusa et al., 2022; Kohlmeier et.al, 2020; Schipper et.al, 2017; Santagata et al., 2007; Stigler & Thomson, 2009).

The effectiveness of professional development can be influenced by inquiry investigations involving self-study (a proxy for lesson study). Inquiry investigations in this context consists of comparing different curriculum and pedagogical techniques, improving lesson plans, discussing student ideas in a specific subject area, and connecting student ideas to instruction. Lotter et al. (2014) designed a two-week teacher summer institute followed by four sessions of lesson study during one academic year for 36 inservice mathematics middle school teachers. The participating teachers in this study reported an increased understanding of content knowledge, inquiry instruction, and instructional strategies resulting from the professional development. The teachers achieved a statistically significant increased score on an observational protocol, entitled the Reformed Teacher Observation Protocol (RTOP), after participating in the teacher institute.

Lewis et al. (2009) also applied a lesson study approach of professional development during a session of a two-week summer workshop for teachers in a large urban school district. In this study, the teachers were involved in a focal lesson study group. In the first phase of lesson study, the teachers studied their state content standards and discussed and solved problems associated with such content. In subsequent phases, the participating teachers selected, observed, and collected student data from a research lesson and then discussed, revised, and re-taught their lesson (as lesson study) to another group of students. The teachers were videotaped as they wrote a lesson plan, taught and observed the research lesson, revised the lesson plan, and re-taught the lesson. Data were also collected for this research from teacher group meetings, student work, observational field-notes, and follow-up conversations among lesson study teachers. Results of this study indicate that disciplinary instruction was improved through lesson study by improving not only teachers' content and pedagogical knowledge, but also their teaching beliefs, their sense of community, and their access to teaching-learning resources.

In a South African case study conducted by Helmbold et al. (2021), an internationally well-established lesson study framework was implemented in the professional development of six mathematics teachers. Teachers' data were collected from a focus group interview, weekly teacher journaling, research questionnaires during mid-study and at the end of the study, and live recordings of teacher professional development meetings. The results of this study indicated that lesson study improved the teachers' pedagogical content knowledge and promoted the diversity of teaching resources for the participating teachers. Lesson study created an opportunity for the personal awareness of certain content knowledge shortcomings on the part of the participating teachers. It opened up the practice of other pedagogical approaches, such as problem-solving, that were not popular teaching approaches among the participating teachers before the study. Lesson study raised the teachers' expectations for their students and fostered higher-level thinking in their classroom. It also helped build teaching efficacy in the teachers, enhanced their content and pedagogical knowledge, and enabled them to integrate cooperation and curriculum development beyond the grades in which they were teaching.

Numerous positive effects of lesson study on teachers' competency were also found in research by Listyani and colleagues (2008). Teacher competency was increased, which included pedagogic competence, professional competence, and social competence. In this particular study, the lesson study activities included planning, implementation and observation, and reflection. The participating teachers discussed measurement tools for student learning, lesson planning, students' work artifacts, media, and evaluation instruments in their lesson study. Each teacher acted as a teaching model twice and observed other members' teaching process for eight rounds through the lesson study. Data were collected from student questionnaires and interviews, teacher interviews, and in-class observations at the end of the lesson study participation. Results of this research demonstrated that more than 80% of students agreed they were more involved in their learning after the teachers completed the lesson study. The teacher participants reported that they had better classroom management and improved lesson design and media development skills resulting from the lesson study experience. It is for these reasons that the present study has incorporated lesson study as a critical component of the researched teacher intervention.

The Effects of Linking Content and Pedagogical Knowledge to Mathematics Teacher Professional Development

While there are many skills, strategies, and understandings one needs to effectively teach mathematics, deep interconnected content knowledge is crucial to impactful teaching in secondary schools (Windschitl, 2009). Scholars argue that mathematics teachers, particularly in middle schools, fall short in their understanding of the contemporary mathematics content they are required to teach (Schmidt et al., 2007). A study by Schmidt and colleagues (2007) compared the preparation of 2,627 pre-service middle school mathematics teachers who were in their last year of training across 34 institutions in six countries. Findings from this mega study confirmed that countries with higher student achievement had teachers with a deeper understanding of the content required for mathematics instruction.

Research, however, has shown that having adequate teacher content knowledge is a necessary yet insufficient condition for quality instruction. Wenglinsky's (2000) study of the relationship between teacher professional development and student achievement revealed that classroom practices, in the form of pedagogical knowledge and teachers' use of higher order thinking skills, is a stronger predictor of students' academic success than a teacher's content knowledge alone (Wenglinsky, 2000). This research demonstrates that both content and pedagogical knowledge applied to content is essential to teacher and student success and, thus, provided a basis for the presented professional development intervention's approach.

Shulman (2004) posited that it is a teacher's ability to recognize how to make content meaningful through disciplinary pedagogy that determines the teacher's skill to make transformative curriculum decisions. Further, Mishra and Koehler (2006) conceptualized a framework for pedagogical content knowledge that describes the complexities of the classroom as dependent on the context of the teachers' everyday realities and the "… thoughtful interweaving of all three key sources of knowledge: technology, pedagogy, and content" (p. 1029). In the intervention for this manuscript, the teachers' pedagogical content knowledge for mathematics was expanded upon. Therefore, it was determined that effective teacher professional development, which was intended to innovate and change the quality of mathematics curriculum, must accentuate both subject matter content knowledge and disciplinary pedagogical knowledge while emphasizing higher order thinking skills within the situational context of a school. This was at the core of the described intervention for this manuscript, in which the school is connected to a homogeneous community context.

One of the research questions Anabousy and Tabach (2022) sought to answer in teacher professional development was whether the teachers' pedagogical technology knowledge would change significantly after participating in a community of inquiry framework focused professional development program. In Anabousy and Tabach's research, math teachers from 12 different schools in Israel participated in a six-month professional development program that focused on pedagogical technology knowledge and a community of inquiry. A pedagogical technology knowledge questionnaire was used in the referenced study to assess teachers' orientation in their beliefs and confidence, pedagogical knowledge, technology instrumental genesis, knowledge of mathematical content, and personal background. The community of inquiry framework encouraged teachers' collaborative practices with a focus on expanding their knowledge in technology integration. Anabousy and Tabach's (2022) results revealed that teachers had higher pedagogical technology knowledge after participating in the prescribed professional development program.

Teachers who are experts in the content they teach have been found to be more effective at producing academically successful students. Disciplinary focused professional development can help improve teacher content knowledge by embedding teachers in environments with focused content and assisting them in effectively teaching such content to their students. Through this type of professional development, teachers can develop disciplinary pedagogical knowledge—in other words, an understanding of how students learn specific content. In Mundry's (2005) research, teachers participated in a professional development experience where they explored mathematics case studies that integrated content and content learning for teachers. Students of the teachers who participated in the program made gains in mathematics test scores whereas the students of the teachers who did not participate demonstrated no significant math test score gains. Therefore, professional development that integrates content and content learning for teachers results in students' achievement gains.

Teaching Mathematics Through Problem Solving to Foster Student Learning

The National Council of Teachers of Mathematics (NCTM) (2000) refers to mathematical problem solving as mathematical tasks designed to provide intellectual challenges for strengthening students' math understanding. There has been growing agreement among math educators that teaching mathematics through problem solving fosters students' learning because it provides opportunities for students to play a very active role in their mathematics learning by exploring real life situations and developing their own strategies to come up with solutions (Suryani et al., 2020; Cai, 2003, 2010; Hiebert et al., 2002; Lesh & Doerr, 2013; Lester, 2013; Schoenfeld, 1994). Such research has revealed that problem-solving approaches enhance the development of the logical thinking aspects of mathematics and contribute to its practical use (Resnick, 1987). For more than two decades, problem solving has played a prominent role in mathematics education for K-12 students.

In a study conducted by Khalid et al. (2020), 172 sixth-grade Malaysian students participated in a creative problem-solving intervention. This particular intervention enabled students and teachers to reason and develop creative solutions through critical questioning strategies. This study explored the relationship between students' creativity (on the Torrance Tests of Creative Thinking) and problem-solving ability. The study's results indicated that students' creativity and problem-solving ability increased after the students engaged in problem solving instruction and practice. The study suggests that the collaborative creative problem-solving approaches used in the study's intervention improved students' skills. It also helped students to enhance their confidence, collaborate with peers, and develop greater potential in learning mathematics.

An essential component of teaching mathematics through problem solving involves a teacher's proficiency in incorporating this into their pedagogy and associated curriculum. However, understanding how to engage in such practice is not necessarily obvious to mathematics teachers. Previous research has revealed that mathematics teachers experience challenges in designing and implementing problem-solving instruction (Goldberg & Bush, 2003).

Effects of Motivation on Student Achievement in Mathematics

Research in secondary STEM education has shown that, although students' cognitive abilities and socio-demographic backgrounds are important predictors of their achievement, affective constructs, such as motivation, have emerged as significant factors affecting students' performance. For example, Singh et al. (2002) examined the effects of motivation, attitude, and academic engagement on students' achievement in mathematics. Results from structural equation models revealed positive effects of students' motivation, attitude, and academic time spent on homework on their achievement. Although the strongest effect was found to be academic time spent on homework, the study also revealed statistically strong effects of motivation and positive attitude on students' academic performance.

In another study, Cleary and Chen (2009) found that achievement groups in seventh grade identified as high, moderate, and low were clearly differentiated across both self-regulation and motivation characteristics. Notably, research has found that students' achievement in mathematics is associated with intrinsic motivation more than with extrinsic motivation. As such, students who are interested in mathematics and enjoy it demonstrate higher achievement. Using the Rasch estimates of the Programme International Student Assessment (PISA) mathematics test scores and questionnaire responses of 107,975 15-year-old students, Chiu and Xihua (2008) compared familial and motivation effects on students' mathematics achievement across 41 countries. They found that in most countries, student achievement was highly associated with intrinsic motivation. The authors concluded that "students learn more if educators focus their efforts on raising students' academic interests rather than emphasizing extrinsic motivation" (2008, p. 332).

In a study conducted by Xiao and Sun (2021), data from the PISA was analyzed to assess students' mathematics, reading, and science competencies, as well as anxiety, motivation, and other learning-related factors. These researchers sought to determine the different combinations of students' subgroups based on their motivational and affective factors and the relationship between these factors and their achievement in mathematics. The study results revealed that students with high motivation and low anxiety performed better on the PISA assessment than all other classes. Although anxiety played a significant role in students' achievement in Xiao and Sun's (2021) research, regardless of their motivation, educators focusing on providing theoretical and practical support to these subgroups of students may increase their motivation and lower their anxiety.

Furthermore, findings from other studies suggest a reciprocal effect between student motivation and teacher behavior and between student motivation and student achievement. In a study that employed a multivariate latent change model, Gottfried et al. (2007) examined the longitudinal relationship between academic intrinsic mathematics motivation and mathematics achievement among participants aged 9-17 years. The results revealed that mathematics achievement was a significant contributor to the developmental decline in students' intrinsic mathematics motivation. In addition, academic intrinsic mathematics motivation was found to be related to initial and later levels of mathematics achievement.

Examining students' motivation is an important endeavor because it has consistently been shown to predict students' performance in mathematics. Students' in-class experiences affect their motivation and engagement, and teachers are in a position to best promote both. Therefore, the professional development intervention described in this manuscript was designed to impact both students' achievement in mathematics and their interest in and motivation for it by deliberately instructing and guiding teachers on how to use specific teaching methods as a means of increasing students' interest in, motivation for, and achievement in mathematics.

METHOD

Data Collection

Study Setting

This research study took place in five middle schools in an urban mid-sized public school district. Three of the middle schools involved in the study are traditional middle schools, thereby serving students in grades 6-8, one school site is a K-8 school, and one is a 4-8 school. Only sixth- through eighth-grade teachers and classrooms in the K-8 and 4-8 schools were a part of the teacher professional development intervention. The school district in which the study was conducted is situated amidst a community that is struggling significantly both economically and resource-wise and, therefore, the educational needs in the school district are profound. All schools that were targeted for this research intervention were in Program Improvement (PI, a federal achievement category) status at the start of the intervention and, accordingly, their academic performance indices were low. In this school district, the students' academic needs were pronounced at the start of this intervention, with 55-82% of the sixth- through eighth-grade students scoring non-proficient (basic to far below basic) on state standards tests (SST).

Study Sample

The study sample for this research included two groups: middle school teachers and their sixth- through eighth-grade students. There was a total of 64 teachers in the study including general education math teachers and a handful of special education math teachers. Principals and other site and district administrators also participated in the intervention but they were not a part of the study sample. The student sample consisted of 5,505 students. The participating students were primarily of Latinx/ Hispanic decent; however, there was ethnic and cultural diversity in the sample. The distribution was 88.3% Latinx/Hispanic, 7.6% African American, 1.4% White, and the remaining 2.7% from other ethnicities. Linguistically, 40.1% of the students in the sample spoke Spanish as a primary language at home and 33.2% of the students were categorized as limited English proficient at the onset of the intervention. Additionally, 84.3% of the students received free or reduced lunch across the five middle schools, thereby indicating socio-economic need on the part of the participating students. The students' academic needs were great prior to the start of the intervention, with 57-75% sixth- through eighth-grade students scoring non-proficient (basic to far below basic) on the SST in math. This achievement gap persisted for six years prior to the intervention and was particularly dismal for ethnic minority student groups (~13% lower than non-minorities) and English learners (~10-23% lower than "English only" students).

Recruitment Process

In terms of recruitment for the study intervention, both the teacher and student sample were recruited through the school district's administration and therefore recruitment was inclusive. Accordingly, all sixth- through eighth-grade math teachers in the district and their students participated in the intervention. Because the data provided by the district was masked for identification for human subject protection (and this decision was made prior to the intervention), a parent-child "opt out" procedure was employed in the recruitment design; no families chose to opt out of the research intervention.

The participating teachers were compensated for participation in the intervention via research funds during the summer teacher academy and via releases with substitute teachers during the academic year, as academic year intervention sessions were conducted during the teachers' contracted day.

Intervention and Differences in Control/Comparison Group Condition

The study intervention was targeted at the middle school teachers in the district with an expressed intent to positively impact students' achievement in mathematics. Therefore, the teacher professional development had seven enabling components that were well established with collaboration between the supporting university (the researchers) and the school district. The program represented a synergy of the components to specifically support the needs of the middle school teachers and their students. The enabling structures included: (a) university level national research laboratories; (b) a teacher training leadership team; (c) a content expert math advisory team; (d) mathematics teacher professional development using a summer teacher academy and associated follow-up; (e) use of lesson study (f) use of diagnostic teaching, inquiry focused mathematics, and data driven decision-making; and (g) foci on integrated mathematics. The teachers received either two or three years of the professional development intervention (the treatment), and therefore their sixth- through eighth-grade students received either a single, double or triple year dose of the intervention depending on the year that they entered middle school. The specific teacher interventions within the professional development consisted of: use of academic language in math contexts, applied mathematics and real-life problem solving, strategies for effective and efficient use of mathematics informational texts, effective uses of technology in the classroom, strategies for integrating math across the curriculum, a teacher summer academy approach to improving teachers' content knowledge (with content experts), and strategies for nimble data driven lesson design focused on inquiry and learning cycles pedagogical structures. The teachers had weeklong summer teaching academies followed by grade level mini camps during the school year at school sites. Lesson study was a major component of the professional development. A typical yearlong "dose" of professional development included 30 hours of the summer teacher academy, followed by 15-20 hours per semester during the academic year (50-60 professional development hours per year =1 dose).

The research included a one-year quasi control condition in the intervention design. Accordingly, the three traditional sixth- through eighth-grade middle schools began implementation in the first intervention year, and the two non-traditional schools (K-8 and 4-8) served as quasi "control" schools during that year and began the professional development intervention the following year. Once the implementation of the full intervention across all five schools was in place, the intervention was delivered in groups and therefore no differences in treatment were noted (except for dose because of the deliberate delayed start of implementation for two of the schools.) There was no randomization of treatment in this study design. The selection of schools and years of implementation was made at the request of the school district's administration because the district was transitioning from junior high school models to middle school models just prior to the start of the intervention.

Research Questions

This study responded to three research questions:

- Research Question 1: What is the impact of lesson study focused, content rich mathematics inservice teacher professional development on middle school teachers' mathematics teaching efficacy?
- Research Question 2: What is the impact of the described teacher professional development on middle school mathematics teachers' students' mathematics achievement?
- Research Question 3: What is the impact of the described teacher professional development on middle school students' mathematics interest and motivation?

Outcomes Measurement and Instrumentation

The outcomes for this research were measured both at the teacher and student levels. Primarily, instrumentation used for the study was standardized, statistically reliable, and highly validated using item response theory (Wilson, 2011). Because the research was a mixed design study, field notes from lesson study group meetings, rubrics and planning forums, were also used and needs assessment (open and close set) questionnaire data was collected to add depth to the breadth of the data in this research.

Instrumentation for Teachers

The following instrumentation was used for teachers:

- Math Teaching Efficacy Beliefs Instrument (MTEBI): This is a Likert-type questionnaire that measures teachers' mathematics teaching efficacy.
- Lesson Study Scoring Rubric: This is a multi-dimensional observational rubric (scaled through full implementation of lesson study), called the Teacher Performance Observational Rubric (TPOR; adapted from PACT; 8 points possible across rubric dimensions).
- Interview Protocol: These were periodic focus group interviews that were conducted with teachers in the intervention to assign voice to their experiences during the intervention.
- Lesson Study Structured Field Notes: These were notes that were taken during the lesson study planning and implementation processes.
- Teacher Feedback and Ongoing Needs Assessment Questionnaire: This was an electronically administered questionnaire that provided formative feedback and needs identification from the participating teachers.

Instrumentation for Students

Instrumentation for the students included the following:

- State Standards Tests (SST) in Mathematics (grades 6-8): These are the state adopted standardized and validated achievement assessments.

- District Benchmark Exams in Mathematics (grades 6-8): These are the target school districts' mathematics benchmark examinations for grades 6-8 that are administered quarterly district wide. They are criterion referenced with district determined "cut scores" as passing score equivalents.
- Motivation for Mathematics Questionnaire: This Likert-type questionnaire was designed, tested, and validated using IRT (Wilson, 2011) by one of the authors and administered by the research team via the teachers as an annual pre-post comparison at the start and end of every academic year during the study. This questionnaire contained 10 subscales associated with motivation, curiosity, engagement and efficacy in mathematics. The participating students received the mathematics motivation questionnaire in their mathematics class.

Instrument Reliability and Validity

The study's instrumentation has been tested for reliability and validity. The TPOR consists of six teacher instructional performance dimensions (each with one or more sub-dimension), including planning, assessment, instruction, reflection, academic language, and accommodations for diverse learners, and its reliability is strong (Cronbach's alpha = .93; Ragusa, 2011). It was modeled after a combination of the Performance Assessment of Teachers and Ball's teacher observational assessment. It was tested for validity and reliability using Wilson's four building blocks of item response theory (Wilson, 2011). Collectively, in terms of statistical power of the full set of instruments, given that there were 5,505 in the student sample, statistical power was robustly achieved (Cohen, 1992). The data for this research was collected, analyzed, stored, and protected in an ethical manner and was fully approved by the University's Human Subjects Review Board at the primary author's university.

Statistical/Analytical Methods

The study included a mixed methods research design. Accordingly, data analyses were both quantitative and qualitative. In addition to conducting descriptive analyses, pre-post comparisons, including *t*-tests with effects sizes (Cohen's *d*), were computed to illustrate the diversity in the study sample. Correlation analyses were conducted as a precursor to multivariate and multilevel hierarchical linear modeling (HLM). Multiple means of model fitting were used for the HLM analyses. SPSS (version 22) and HLM (version 7.1) software were used for these quantitative analyses.

For the qualitative components of the research, categorization of teacher data with frequency distributions was conducted. Specifically, open-ended responses to the teacher questionnaires (needs assessment and evaluative feedback from professional development), lesson study and planning observations, and field notes were analyzed using well established thematically focused qualitative analyses. NVivo (version 10) was used for these analyses. The data were coded and thematically categorized using a constant, comparative method (Lincoln & Guba, 1985). Special attention was paid to disconfirming evidence and outliers in data coding, as well as elements of frequency, extensiveness, and intensity within the data. Ideas or phenomena were initially identified and flagged to generate a list of internally consistent, discrete categories, followed by fractured and reassembled (axial coding) categories by making connections between categories and subcategories to reflect emerging themes and patterns in the data. Categories were then integrated to form grounded theory and aligned with existing teacher development theory using selective categorization to clarify concepts and to allow for response interpretations and conclusions associated with the teachers' perceptions of success and challenges of the lesson study, their professional development, and their students' successes and challenges. Frequency distributions of the coded and categorized data was computed. The intent of this intensive qualitative analysis was to identify patterns, make comparisons, and contrast one teacher's or groups' discussion, action, and voice with another throughout the study.

Limitations of Data and Analytical Methods

The data for this research was limited by a few constraints. First, all student-level data was collected by the participating classroom teachers. Therefore, there was missing data that was accounted for by reliable and widely accepted statistical procedures for managing such data. In terms of an analytical approach, because multivariate approaches and multi-level hierarchical linear modeling were used, these limitations were not profound. The primary limit analytically emerged when, in an attempt to build and fit a three-level model (students nested in teachers' classrooms and nested in schools), it was determined that due to the shared and homogenous nature of the five schools' population characteristics, no statistically significant effects were noted at Level III of the model, therefore causing a need to return to a two-level hierarchical model. The interpretation of this analytical structure is described in the results below.

RESULTS AND DISCUSSION

Estimates of the Intervention's Effect on All Outcomes (With Subgroups)

Given that this research involved teachers and their impact on students, the researchers wanted to measure the impact of the teachers' intervention—specifically their change in instructional practice and knowledge—on their students' achievement, motivation, and interest. Accordingly, the results are divided by study population (below) and then the combined/interactive effects are described and illustrated. Within the teacher-related results, both quantitative and qualitative results are presented.

Teacher Effects

Teacher effects resulting from the teacher professional development intervention relate both to teaching efficacy (measured by the MTEBI) and teacher performance (measured by an observational rubric, the TPOR). These results are interesting and diverse.

All teachers participated in lesson study and such lesson study consisted of preparation of a collaborative lesson plan, assessment planning, a videotaped teaching event, debriefing and scoring sessions, and plan revisions. The TPOR was scored multi-dimensionally on an 8-point rubric with a score of eight being the highest possible score. The mean TPOR scores by subgroup of teachers is provided in Table 1.

The teachers' teaching events were scored only one time during the intervention and, therefore, there was no pre-post intervention comparison in the study results. With regard to teachers' efficacy for teaching math, moderate difference between scores for the three years/grades were indicated: Grade 6 Teachers: $M_{pre} = 3.03$, $SD = .231$, $M_{post} = 3.62$, $SD = .418$ $t(22) = 6.99$; Cohen's $d = 1.723$; Grade 7 Teachers: $M_{pre} = 3.17$, SD = .224, $M_{post} = 3.82$, $SD = .222$ t(29) = 10.3; Cohen's $d = 2.355$; Grade 8 Teachers: $\mathrm{M_{pre}} = 3.19$, SD= .246, $\mathrm{M_{post}} = 3.75$, $SD = .380$ $t(32) = 10.7$; Cohen's $d = 1.751$. There were also considerable variations in results across teacher subgroups. Specifically, teachers with less than two years of teaching experience significantly improved (mean difference = .49, $t(5)= -3.726$; $p < .05$) over time. Teachers who taught only seventh grade improved at a somewhat lower yet still a significant rate (mean difference= .30, $t(6) = -5.81$; $p < .001$). Teachers who taught only one grade had a significant increase in teaching efficacy (mean difference =.20, $t(15) = -3.704$; $p <.01$) across time. Additionally, when correlated with the Teacher Performance Observational Rubric (TPOR),

Table 1. TPOR Scores by Grade (8-Point Rubric Score)

Grade 6 Mean = 6.21 *SD* = 0.88	Grade 7 Mean = 5.93 *SD* = 1.26	Grade 8 Mean = 6.04 *SD* = 1.29

with eight points possible, the scores were highly correlated with teachers' teaching efficacy at the end of the intervention ($M = 5.47$, $SD = 11.03$; $r = .47$).

Qualitatively, the teachers' perception of the professional development, their reactions, and their progress were remarkable as well. Results are illustrated below in Table 2 as frequency distributions categorically.

The combined teaching efficacy and results indicate that both teacher attitude toward teaching and teaching efficacy—in other words, their confidence specific to teaching math—increased over time. Additionally, the teachers became more engaged and reported being more knowledgeable about teaching practices.

Student Effects

The participating students' results were comprehensive, diverse, and impactful. With regard to changes in student achievement, while the results of the students' SST in mathematics were variable, on the whole, all five schools had gains between the pre intervention period and last year of the intervention. The same was true when tracking the trajectories of the schools' Academic Performance Index (API).

As illustrated by Figures 1 and 2, students in grades 6-8 in the sample experienced steady gains in mathematics standardized testing across the three intervention years. The test score growth varied across years for students. In the baseline year, three of five schools' percentage of students scoring proficient or advanced in mathematics was below the district averages across grades. Between the first and second intervention years, the intervention schools had an 8.9% growth in percentage of students scoring at the proficient or advanced level in standardized mathematics testing and the quasi-control schools (two schools) had only 7.05% mathematics growth in terms of percentage of students scoring proficient or advanced. Between the first and last years of intervention, while there was variability in growth, the percent of growth in students scoring proficient or advanced in standardized mathematics testing was 12.58%, which was above the district growth percentage.

Motivation for Mathematics

With regard to mathematics motivation, the study results indicated that there were increases in intrinsic motivation for the students across intervention years, with some variability across sub-constructs of motivation. Specifically, over time, there were larger gains in intrinsic motivation and

Table 2. Teacher Perceptions of Professional Development (PD) Intervention Across Time

Coded Category	Pre- PD Freq. (%)	During PD Freq. (%)	Example
Teaching Enthusiasm	49 (44.1)	63 (36.2)	"I was beginning to feel very burnt out." Now I'm feeling invigorated and raring to go in my teaching. I look forward to the days."
Pedagogical Language	17 (15.3)	42 (24.1)	"I find using ACT work best."
Collaboration	22 (19.8)	27 (15.5)	"Lesson study gives us time to work across schools. We never had this before CPEC."
Help Seeking	19 (17.1)	13 (7.5)	"I am glad we are able to keep (Name, math coach). I rely on her daily!"
Help Provision	4 (3.6)	29 (16.7)	"I don't care if you have a look at my video. I'll bring it by along with some materials and we can plan the next couple of weeks."
TOTAL:	**111 (100)**	**174(100)**	

Figure 1. Cross-School Comparison of Academic Performance Index (API)

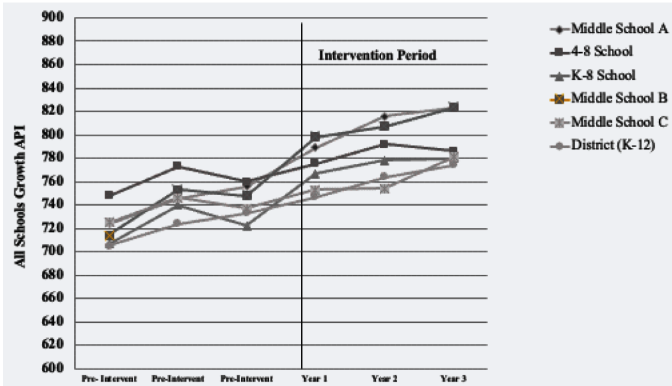
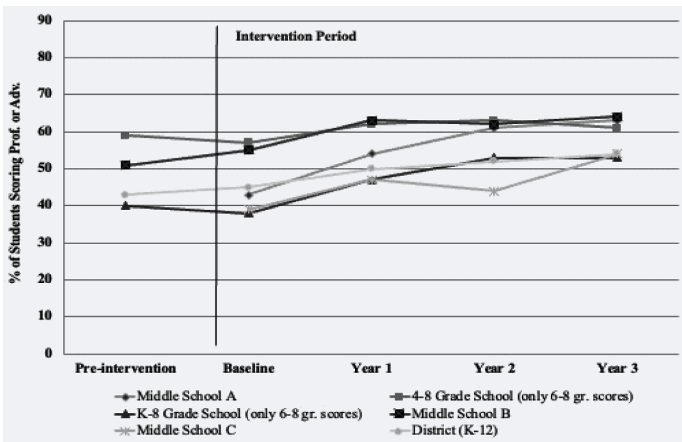


Figure 2. Cross-School Comparison-Student Achievement- Math



decreases in extrinsic motivation. In other words, the students in the sample improved their ability to be motivated to engage in math without external rewards of any type, which is supported by the literature as leading to improved learning and eventual achievement (Guthrie & Klauda, 2014). For example, the motivational sub-construct of math enjoyment was significantly increased pre-to-post intervention ($t(2315) = 3.45$; $p < .001$; Cohen's $d = .52$), as was the participating students' belief in math importance ($t(2315) = 4.68$; $p < .001$; Cohen's $d = .57$).

Linear regression analyses of mathematics motivation using the students' standardized math achievement scores as a dependent variable were conducted. These results indicated the following.

Math Motivation predicted student achievement in Math across grade levels.

- Gr. 6: $R^2 = 0.24$, $F(10, 3017) = 94.96$, $p < .001$
- Gr 7: $R^2 = 0.17$, $F(10, 2894) = 59.56$, $p < .001$
- Gr. 8: $R^2 = 0.16$, $F(10, 3028) = 59.50$, $p < .001$

This result is also supported by the motivation and achievement research, conducted by Guthrie and Klauda (2014).

Combined Integrative Effects

After completing analyses for both the participating teachers and their students and because the primary goal of our research project was to test the impact of a teacher professional development intervention on student achievement, hierarchical linear modeling was used for analyses. The research team proceeded through various iterations of model fitting following well-established research conducted by HLM founders Raudenbush and Bryk (2002). Accordingly, a three-level model with students nested in teachers' classrooms, which were nested in schools was utilized initially. After testing this model (and its null counterpart,) it was determined that there were no statistically significant school effects on the first two levels of the model (*interceopt1/intercept2: Chi Sq. = .01726, p >500*) and, therefore, the analysis shifted to a two-level hierarchical linear model with students nested in their teachers' classrooms. As such, the models in statistical notation and their associated complete results are described in Tables 3 through 6. The model was run with outcome (dependent) variables for math benchmark scores and math standardized testing for each of the three grade levels (6-8). All variables for each model are labeled within each model (see complete variable listing in Appendix A). Both unconditional and conditional models are indicated below with model equations in numbered order. Dummy coding was used for variables including teacher credentials (whether the participating teacher had a single subject credential), gender, and ethnicity. Both aggregate and non-aggregates of predictor variables were utilized in the models, and group centering was utilized and is clearly delineated in each model. The summary results are described in narrative form below Table 6. A full listing of abbreviations from the models is included as Appendix B.

HLM Models (Mathematics Benchmark as Dependent Variable)

Full unconditional model

Student-level model:

$$MATH\ BENCHMARK_{ij} = \beta_{0j} + r_{ij} \quad (1)$$

Table 3. Unconditional Model using Math Benchmark as Dependent Variable

Model	Fixed Effects	Coeff.	T-Ratio	P	Reliability
Grade 6 Model	Intercept 1, β_0				0.860
	Intercept 1, γ_{00}	0.667	54.840	.000	
Grade 7 Model	Intercept 1, β_0				0.866
	Intercept 1, γ_{00}	0.696	59.181	.000	
Grade 8 Model	Intercept 1, β_0				0.946
	Intercept 1, γ_{00}	0.457	15.139	.000	

Table 4. Conditional Model Using Math Benchmark as Dependent Variable

Model	Fixed Effects	With Aggregate			Without Aggregate		
		Coeff.	T-Ratio	P	Coeff.	T -Ratio	P
Grade 6 Model	Model for Mean Benchmark β_0						
	Intercept (γ_{00})	0.537	1.406	.182	0.784	6.399	.000
	TEACHER DOSE (γ_{01})	-0.011	-0.731	.477	-0.003	-0.144	.887
	TEACHER PERFORMANCE (γ_{02})	0.007	0.492	0.630	0.004	0.250	.806
	TEACHER EFFICACY (γ_{03})	0.008	0.283	0.781	0.041	1.291	.214
	TEACHER SINGLE CREDENTIAL (γ_{04})	-0.005	0.231	.821	-0.008	-0.294	.773
	TEACHER EXPERIENCE (γ_{05})	-0.022	-1.885	.080	0.021	-1.454	.164
	MATH INTRINSIC MOTIVATION MEAN (γ_{06})	2.654	2.934	.011	-	-	-
	MATH EXTRINSIC MOTIVATION MEAN (γ_{07})	1.926	2.624	.020	-	-	-
	MATH MOTIVATION MEAN (γ_{08})	-4.594	-2.942	.011	-	-	-
	Model for Math Intrinsic Motivation Slope β_1						
	Intercept (γ_{10})	0.416	13.119	.000	0.416	13.121	.000
	Model for Math Extrinsic Motivation Slope β_2						
	Intercept (γ_{20})	0.438	12.496	.000	0.438	12.501	.000
	Model for Math Motivation Slope β_3						
	Intercept (γ_{30})	-0.736	-11.545	.000	-0.736	-11.549	.000
	Model for GENDER Slope β_4						
	Intercept (γ_{40})	-0.019	-3.255	.001	-0.019	-3.240	.001
	Model for HISPANIC Slope β_5						
	Intercept (γ_{50})	-0.027	-1.201	.230	-0.027	-1.189	.235
	Model for AFRICAN AMERICAN Slope β_6 .						

continued on following page

Table 4. Continued

Grade 6 Model	Intercept (γ_{60})	-0.097	-3.895	.000	-0.097	-3.886	.000
	Model for WHITE Slope β_7						
	Intercept (γ_{70})	-0.013	-0.372	.710	-0.013	-0.357	.721
	Model for FILIPINO Slope β_8						
	Intercept (γ_{80})	0.038	0.965	.334	0.038	0.974	.330
	Model for STUDENT DOSE Slope β_9						
	Intercept (γ_{90})	0.028	-2.098	.036	-0.02	-2.011	.044
Grade 7 Model	Model for Mean Benchmark β_0						
	Intercept (γ_{00})	0.058	0.128	.900	0.767	5.215	.000
	TEACHER DOSE (γ_{01})	-0.005	-0.218	.831	-0.012	-0.549	.590
	TEACHER PERFORMANCE (γ_{02})	-0.002	-0.113	.911	-0.002	-0.135	.894
	TEACHER EFFICACY (γ_{03})	-0.009	-0.163	.872	0.027	0.506	.619
	TEACHER SINGLE CREDENTIAL (γ_{04})	-0.018	-0.547	.593	-0.007	-0.196	.847
	TEACHER EXPERIENCE (γ_{05})	-0.002	-0.110	.914	-0.009	-0.550	.590
	MATH INTRINSIC MOTIVATION MEAN (γ_{06})	0.590	0.435	.670	-	-	-
	MATH EXTRINSIC MOTIVATION MEAN (γ_{07})	0.644	0.694	.499	-	-	-
	MATH MOTIVATION MEAN (γ_{08})	-1.002	-0.451	.659	-	-	-
	Model for Math Intrinsic Motivation Slope β_1						
	Intercept (γ_{10})	0.318	11.040	.000	0.318	11.044	.000
	Model for Math Extrinsic Motivation Slope β_2						
	Intercept (γ_{20})	0.315	9.987	.000	0.316	9.991	.000
	Model for Math Motivation Slope β_3						
	Intercept (γ_{30})	-0.525	-9.113	.000	0.526	-9.117	.000
	Model for GENDER Slope β_4						

continued on following page

Table 4. Continued

Grade 7 Model	Intercept (γ_{40})	-0.017	-3.039	.002	-0.017	-3.036	.002
	Model for HISPANIC Slope β_5						
	Intercept (γ_{50})	-0.022	-0.983	.326	-0.022	-0.973	.331
	Model for AFRICAN AMERICAN Slope β_6						
	Intercept (γ_{60})	-0.073	-2.956	.003	0.073	-2.944	.003
	Model for WHITE Slope β_7						
	Intercept (γ_{70})	0.065	-1.991	.047	-0.066	-2.040	.041
	Model for FILIPINO Slope β_8						
	Intercept (γ_{80})	0.075	2.090	.037	0.076	2.101	.036
	Model for STUDENT DOSE Slope β_9						
	Intercept (γ_{90})	0.018	3.231	.001	0.018	3.252	.001
Grade 8 Model	Model for Mean Benchmark β_0						
	Intercept (γ_{00})	1.962	2.457	.028	0.401	2.261	.037
	TEACHER DOSE (γ_{01})	0.106	2.842	.013	0.082	2.033	.058
	TEACHER PERFORMANCE (γ_{02})	0.014	0.563	.583	0.009	0.317	.755
	TEACHER EFFICACY (γ_{03})	0.160	1.364	.194	-0.100	-0.899	.381
	TEACHER SINGLE CREDENTIAL (γ_{04})	-0.074	-1.141	.273	-0.104	-1.423	.173
	TEACHER EXPERIENCE (γ_{05})	-0.049	-1.275	.223	0.012	0.281	.782
	MATH INTRINSIC MOTIVATION MEAN (γ_{06})	-6.812	-3.212	.006	-	-	-
	MATH EXTRINSIC MOTIVATION MEAN (γ_{07})	4.151	-2.228	.043	-	-	-
	MATH MOTIVATION MEAN (γ_{08})	10.561	2.811	.014	-	-	-
	Model for Math Intrinsic Motivation Slope β_1						
	Intercept (γ_{10})	0.005	0.114	.909	0.006	0.117	.907

continued on following page

Table 4. Continued

Grade 8 Model	Model for Math Extrinsic Motivation Slope β_2						
	Intercept (γ_{20})	0.030	0.599	.549	0.031	0.603	.546
	Model for Math Motivation Slope β_3						
	Intercept (γ_{30})	0.010	0.110	.912	0.010	0.106	.915
	Model for GENDER Slope β_4						
	Intercept (γ_{40})	-0.006	-0.627	.531	-0.006	-0.620	.535
	Model for HISPANIC Slope β_5						
	Intercept (γ_{50})	0.020	0.527	.598	0.020	0.543	.587
	Model for AFRICAN AMERICAN Slope β_6						
	Intercept (γ_{60})	0.062	1.482	.138	0.062	1.496	.135
	Model for WHITE Slope β_7						
	Intercept (γ_{70})	0.029	0.527	.598	0.029	0.536	.592
	Model for FILIPINO Slope β_8						
	Intercept (γ_{80})	0.067	1.078	.281	0.067	1.076	.282
	Model for STUDENT DOSE Slope β_9						
	Intercept (γ_{90})	-0.101	-12.925	.000	-0.101	-12.870	.000

MODEL	Random Effects	With Aggregate			Without Aggregate		
		Variance	df	Chi-Square (p)	Variance	df	Chi-Square (p)
Grade 6 Model	Intercept 1, u_0	0.002	14	142.510 ($p <$ .000)	0.003	17	319.248 ($p <$ .000)
	Level -1, r	0.026			0.026		
Grade 7 Model	Intercept 1, u_0	0.003	14	173.683 ($p <$ .000)	0.004	17	250.021 ($p <$ .000)
	Level -1, r	0.022			0.022		
Grade 8 Model	Intercept 1, u_0	0.014	14	386.173 ($p <$ .000)	0.023	17	646.601 ($p <$ .000)
	Level -1, r	0.068			0.068		

Table 5. Unconditional Model Using Math SST as Dependent Variable

Model	Fixed Effects	Coefficient	T-Ratio	P	Reliability
Grade 6 Model	Intercept 1, β_0				
	Intercept 1, γ_{00}	3.432	58.754	.000	0.803
Grade 7 Model	Intercept 1, β_0				
	Intercept 1, γ_{00}	3.696	88.905	.000	0.639
Grade 8 Model	Intercept 1, β_0				
	Intercept 1, γ_{00}	3.464	32.470	.000	0.950

Table 6. Conditional Model Using Math SST as Dependent Variable

Model	Fixed Effects	With Aggregate			Without Aggregate		
		Coeff.	T-Ratio	P	Coeff.	T-Ratio	P
Grade 6 Model	Model for Mean MATH SST β_0						
	Intercept (γ_{00})	4.920	2.589	.021	3.886	6.599	.000
	TEACHER DOSE (γ_{01})	-0.036	-0.464	.650	0.017	.183	0.857
	TEACHER PERFORMANCE (γ_{02})	0.116	1.624	.127	0.086	1.068	.300
	TEACHER EFFICACY (γ_{03})	0.083	0.570	.577	0.098	0.656	.521
	TEACHER SINGLE CREDENTIAL (γ_{04})	-0.165	-1.475	.162	-0.132	-1.006	.329
	TEACHER EXPERIENCE (γ_{05})	-0.137	-2.437	.029	-0.123	-1.828	.085
	MATH INTRINSIC MOTIVATION MEAN (γ_{06})	11.688	2.603	.021	-	-	-
	MATH EXTRINSIC MOTIVATION MEAN (γ_{07})	10.493	2.878	.012	-	-	-
	MATH MOTIVATION MEAN (γ_{08})	-23.107	-2.962	.010	-	-	-
	Model for Math Intrinsic Motivation Slope β_1						
	Intercept (γ_{10})	2.281	11.809	.000	2.281	11.812	.000
	Model for Math Extrinsic Motivation Slope β_2						

continued on following page

Table 6. Continued

Grade 6 Model	Intercept (γ_{20})	2.368	11.104	.000	2.369	11.108	.000
	Model for Math Motivation Slope β_3						
	Intercept (γ_{30})	-4.056	-10.454	.000	-4.057	-10.458	.000
	Model for GENDER Slope β_4						
	Intercept (γ_{40})	-0.092	-2.527	.012	0.091	-2.498	.013
	Model for HISPANIC Slope β_5						
	Intercept (γ_{50})	-0.213	-1.563	.118	-0.213	-1.559	.119
	Model for AFRICAN AMERICAN Slope β_6						
	Intercept (γ_{60})	-0.572	-3.775	.000	-0.570	-3.762	.000
	Model for WHITE Slope β_7						
	Intercept (γ_{70})	0.060	0.275	.783	0.063	0.290	.772
	Model for FILIPINO Slope β_8						
	Intercept (γ_{80})	0.135	0.561	.575	0.144	0.600	.549
	Model for STUDENT DOSE Slope β_9						
	Intercept (γ_{90})	-0.247	-3.041	.002	-0.247	-3.047	.002
Grade7 Model	Model for Mean MATH SST β_0						
	Intercept (γ_{00})	3.054	2.215	.044	3.945	8.839	.000
	TEACHER DOSE (γ_{01})	0.049	0.761	.459	0.022	0.312	.759
	TEACHER PERFORMANCE (γ_{02})	-0.010	-0.221	.828	-0.024	-0.465	.648
	TEACHER EFFICACY (γ_{03})	0.125	0.860	.404	0.129	0.835	.415
	TEACHER SINGLE CREDENTIAL (γ_{04})	0.050	.595	.561	0.124	1.219	.239
	TEACHER EXPERIENCE (γ_{05})	-0.002	-0.054	.957	-0.024	-0.534	.600
	MATH INTRINSIC MOTIVATION MEAN (γ_{06})	7.004	1.925	.075	-	-	-
	MATH EXTRINSIC MOTIVATION MEAN (γ_{07})	3.914	1.417	.178	-	-	-

continued on following page

Table 6. Continued

Grade7 Model	MATH MOTIVATION MEAN (γ_{08})	-10.829	-1.760	.100	-	-	-
	Model for Math Intrinsic Motivation Slope β_1						
	Intercept (γ_{10})	1.880	9.340	.000	1.881	9.346	.000
	Model for Math Extrinsic Motivation Slope β_2						
	Intercept (γ_{20})	1.936	8.777	.000	1.937	8.783	.000
	Model for Math Motivation Slope β_3						
	Intercept (γ_{30})	-3.308	-8.211	.000	-3.310	-8.218	.000
	Model for GENDER Slope β_4						
	Intercept (γ_{40})	-0.067	-1.714	.087	-0.067	-1.703	.089
	Model for HISPANIC Slope β_5						
	Intercept (γ_{50})	-0.091	-0.575	.565	-0.085	-0.538	.590
	Model for AFRICAN AMERICAN Slope β_6						
	Intercept (γ_{60})	-0.317	-1.830	.067	-0.309	-1.785	.074
	Model for WHITE Slope β_7						
	Intercept (γ_{70})	-0.058	-0.254	.799	-0.064	-0.282	.778
	Model for FILIPINO Slope β_8						
	Intercept (γ_{80})	0.538	2.139	.033	0.550	2.187	.029
	Model for STUDENT DOSE Slope β_9						
	Intercept (γ_{90})	-0.038	-1.006	.314	-0.032	-0.847	.397
Grade8 Model	Model for Mean MATH SST β_0						
	Intercept (γ_{00})	-1.815	-0.695	.499	3.336	6.195	.000
	TEACHER DOSE (γ_{01})	0.080	0.656	.522	0.146	1.199	.247
	TEACHER PERFORMANCE (γ_{02})	0.040	0.505	.621	0.022	0.261	.797
	TEACHER EFFICACY (γ_{03})	-0.019	-0.051	.960	0.547	1.619	.124

continued on following page

Table 6. Continued

Grade8 Model	TEACHER SINGLE CREDENTIAL (γ_{04})	0.028	0.131	.898	0.213	0.966	.348
	TEACHER EXPERIENCE (γ_{05})	-0.0002	-0.002	.999	-0.137	-1.097	.288
	MATH INTRINSIC MOTIVATION MEAN (γ_{06})	14.071	2.028	.062	-	-	-
	MATH EXTRINSIC MOTIVATION MEAN (γ_{07})	13.204	2.166	.048	-	-	-
	MATH MOTIVATION MEAN (γ_{08})	-25.938	-2.110	.053	-	-	-
	Model for Math Intrinsic Motivation Slope β_1						
	Intercept (γ_{10})	1.577	10.026	.000	1.576	10.021	.000
	Model for Math Extrinsic Motivation Slope β_2						
	Intercept (γ_{20})	1.648	9.784	.000	1.647	9.777	.000
	Model for Math Motivation Slope β_3						
	Intercept (γ_{30})	2.815	-9.064	.000	-2.813	-9.057	.000
	Model for GENDER Slope β_4						
	Intercept (γ_{40})	-0.150	-4.627	.000	-0.150	-4.625	.000
	Model for HISPANIC Slope β_5						
	Intercept (γ_{50})	-0.061	-0.490	.624	-0.062	-0.501	.617
	Model for AFRICAN AMERICAN Slope β_6						
	Intercept (γ_{60})	-0.291	-2.112	.035	0.293	-2.130	.033
	Model for WHITE Slope β_7						
	Intercept (γ_{70})	-0.097	-0.539	.590	-0.100	-0.552	.581
	Model for FILIPINO Slope β_8						
	Intercept (γ_{80})	0.371	1.791	.073	0.371	1.794	.073
	Model for STUDENT DOSE Slope β_9						
	Intercept (γ_{90})	0.055	2.132	.033	0.053	2.058	.040

continued on following page

Table 6. Continued

MODEL	Random Effects	With Aggregate			Without Aggregate		
		Variance	df	Chi-Square (p)	Variance	df	Chi-Square (p)
Grade 6 Model	Intercept 1, u_0	0.042	14	86.155 ($p <$ .001)	0.066	17	166.697 ($p <$ .001)
	Level -1, r	0.974			0.974		
Grade 7 Model	Intercept 1, u_0	0.010	14	28.319 ($p <$ .01)	0.023	17	52.945 ($p <$ .001)
	Level -1, r	1.085			1.085		
Grade 8 Model	Intercept 1, u_0	0.154	14	351.074 ($p <$ .001)	0.208	17	605.979 ($p <$ .001)
	Level -1, r	0.752			0.752		

Teacher-level model:

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (2)$$

Conditional Models (Benchmarks as Dependent Variable)

Student-level model:

$$MATH\ BENCHMARK_{ij} = \beta_{0j} + \beta_{1j} \times \left(MATH\ INTRINSIC\ MOTIVATION\right)_{ij} + \beta_{2j} \times \left(MATH\ EXTRINSIC\ MOTIVATION\right)_{ij} + \beta_{3j} \times \left(MATH\ MOTIVATION\right)_{ij} + \beta_{4j} \times \left(GENDER\right)_{ij} + \beta_{5j} \times \left(HISPANIC\right)_{ij} + \beta_{6j} \times \left(AFRIC - AMER\right)_{ij} + \beta_{7j} \times \left(WHITE\right)_{ij} + \beta_{8j} \times \left(FILIPINO\right)_{ij} + \beta_{9j} \times \left(STUDENT\ DOSE\right)_{ij} + r_{ij} \quad (3)$$

Teacher-level model with the aggregate of variables:

$$\beta_0 = \gamma_{00} + \gamma_{01} \times \left(TEACHER\ DOSE\right)_j + \gamma_{02} \times \left(TEACHER\ PERFORMANCE\right)_j + \gamma_{03} \times \left(TEACHER\ EFFICACY\right)_j + \gamma_{04} \times \left(TEACHER\ SINGLE\ CREDENTIAL\right)_j + \gamma_{05} \times \left(TEACHER\ EXPRIENCE\right)_j + \gamma_{06} \times \left(\overline{MATH\ INTRINSIC\ MOTIVATION}\right)_j + \gamma_{07} \times \left(\overline{MATH\ EXTRINSIC\ MOTIVATION}\right)_j + \gamma_{08} \times \left(\overline{MATH\ MOTIVATION}\right)_j + u_{0j} \quad (4)$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60}$$

$$\beta_{7j} = \gamma_{70}$$

$$\beta_{8j} = \gamma_{80}$$

$$\beta_{9j} = \gamma_{90}$$

Teacher-level model without the aggregate of variables:

$$\beta_0 = \gamma_{00} + \gamma_{01} \times \left(TEACHER\,DOSE\right)_j + \gamma_{02} \times \left(TEACHER\,PERFORMANCE\right)_j + \gamma_{03} \times \left(TEACHER\,EFFICACY\right)_j + \gamma_{04} \times \left(TEACHER\,SINGLE\,CREDENTIAL\right)_j + \gamma_{05} \times \left(TEACHER\,EXPRIENCE\right)_j + u_{0j} \quad (5)$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{9j} = \gamma_{90}$$

MATH INTRINSIC MOTIVATION, MATH EXTRINSIC MOTIVATION and MATH MOTIVATION have been centered around the group mean.

TEACHER EFFICACY has been centered around the grand mean.

HLM Models (Mathematics SST as Dependent Variable)

Full unconditional model

Student-level model:

$$MATH\,SST_{ij} = \beta_{0j} + r_{ij} \quad (6)$$

Teacher-level model:

$$\beta_{0j} = \gamma_{00} + u_{0j} \tag{7}$$

Conditional Models (SST as Dependent Variable)

Student-level model:

$$\begin{aligned}
&MATH\,SST_{ij} = \beta_{0j} + \beta_{1j} \times \left(MATH\,INTRINSIC\,MOTIVATION\right)_{ij} + \\
&\beta_{2j} \times \left(MATH\,EXTRINSIC\,MOTIVATION\right)_{ij} + \beta_{3j} \times \left(MATH\,MOTIVATION\right)_{ij} \\
&\beta_{4j} \times \left(GENDER\right)_{ij} + \beta_{5j} \times \left(HISPANIC\right)_{ij} + \beta_{6j} \times \left(AFRIC - AMER\right)_{ij} + \\
&\beta_{7j} \times \left(WHITE\right)_{ij} + \beta_{8j} \times \left(FILIPINO\right)_{ij} + \beta_{9j} \times \left(STUDENT\,DOSE\right)_{ij} + r_{ij}
\end{aligned} \tag{8}$$

Teacher-level model with the aggregate of variables:

$$\begin{aligned}
&\beta_{0} = \gamma_{00} + \gamma_{01} \times \left(TEACHER\,DOSE\right)_{j} + \gamma_{02} \\
&\times \left(TEACHER\,PERFORMANCE\right)_{j} + \gamma_{03} \times \left(TEACHER\,EFFICACY\right)_{j} \\
&+ \gamma_{04} \times \left(TEACHER\,SINGLE\,CREDENTIAL\right)_{j} + \gamma_{05} \\
&\times \left(TEACHER\,EXPRIENCE\right)_{j} + \gamma_{06} \times \left(\overline{MATH\,INTRINSIC\,MOTIVATION}\right)_{j} \\
&+ \gamma_{07} \times \left(\overline{MATH\,EXTRINSIC\,MOTIVATION}\right)_{j} + \gamma_{08} \\
&\times \left(\overline{MATH\,MOTIVATION}\right)_{j} + u_{0j}
\end{aligned} \tag{9}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60}$$

$$\beta_{7j} = \gamma_{70}$$

$$\beta_{8j} = \gamma_{80}$$

$$\beta_{9j} = \gamma_{90}$$

Teacher-level model without the aggregate of variables:

$$\beta_0 = \gamma_{00} + \gamma_{01} \times \left(TEACHER DOSE\right)_j + \gamma_{02} \times \left(TEACHER PERFORMANCE\right)_j + \gamma_{03} \times \left(TEACHER EFFICACY\right)_j + \gamma_{04} \times \left(TEACHER SINGLE CREDENTIAL\right)_j + \gamma_{05} \times \left(TEACHER EXPRIENCE\right)_j + u_{0j} \quad (10)$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{9j} = \gamma_{90}$$

MATH INTRINSIC MOTIVATION, MATH EXTRINSIC MOTIVATION and MATH MOTIVATION have been centered around the group mean.

TEACHER EFFICACY has been centered around the grand mean.

Results of the hierarchical linear modeling indicate some interesting effects. For sixth graders in mathematics (using the math outcome variable of math benchmarks), the HLM model results indicate that both intrinsic motivation *(* γ_{10} *= .416, t(13.119); p < .001)* and extrinsic motivation *(* γ_{20} *= .438, t(12.496); p < .001)* had a highly positive effect on student achievement. Sixth-grade student doses of the intervention (via being in participating teachers' classroom) had a moderate negative effect on student achievement *(* γ_{90} *= .028 t(-2.089); p < .05)*. Again, sixth-grade students had a smaller intervention "dose" in duration than any other grade in the intervention as they were only in the intervention for one year before the study's end.

For seventh graders in mathematics (using the benchmarks as an outcome variable/ achievement measure), both intrinsic and extrinsic motivation were found to have a highly positive effect on math achievement, respectively (γ_{10} *= .318, t(11.040); p < .001*), (γ_{20} *= .31, t(9.987); < .001*). The students' dose of the math intervention also had a highly positive effect on their math achievement, (γ_{90} *= .018, t(3.231); p < .001*). For eighth graders in math, using the benchmarks as an outcome variable/achievement measure, the effects were not positively associated with any predictor variables in the model.

When using mathematics standardized scores as the outcome variable, results of the sixth-grade HLM model indicate that both intrinsic and extrinsic motivation were found to have a highly positive effect on students' math achievement, respectively (γ_{10} *= 2.28,1 t(11.809); p < .001*), (γ_{20} *= 2.368, t(11.104); p < .001*). The students' dose of the intervention had a negative effect on their math achievement, (γ_{90} *= -.24, t(-3.041); p < .05*). These effects mirror that of the math benchmark HLM model for sixth graders.

When using math standardized scores as the outcome variable, results of the seventh-grade HLM model indicate that both intrinsic and extrinsic motivation were found to have a highly positive effect on students' math achievement, respectively (γ_{10} *= 1.880, t(9.340); p < .001*), (γ_{20} *= 1.936, t(8.777);*

$p < .001$). These motivational effects mirror that of the math benchmark outcome variable for seventh graders.

When using math standardized scores as the outcome variable, results of the eighth-grade HLM model indicate that both intrinsic and extrinsic motivation were found to have a highly positive effect on students' math achievement, respectively ($\gamma_{10} = 1.57, t(10.026); p < .001$), ($\gamma_{20} = 1.648\ t(9.784); p < .001$). The students' dose of the intervention had a positive effect on their math achievement, ($\gamma_{90} = .05,5\ t(2.132); p < .05$). These effects mirror that of the math benchmark outcome variable for eighth graders.

It is especially noteworthy that most of the teacher factors (independent variables in Level 2 of the model) did not significantly affect their students' achievement in math. The participating teachers' years of experience and teachers' single-subject credential (a proxy for pre-intervention math content knowledge) did not have a significant effect on students' achievement in mathematics.

In summary, the participating students' motivation for math had a positive effect on their achievement. Furthermore, the dose of the students' exposure to the teachers' professional development had a positive effect on students' achievement. The students' socio-demographic factors did not have an effect on their math achievement.

CONCLUSION

In conclusion, this comprehensive teacher professional development intervention was found to be highly effective in positively impacting student achievement, motivation, and interest in mathematics. The findings from this research indicate that both teacher pre-intervention factors and school-level factors did not significantly impact student achievement in mathematics. This supports the understanding that the math professional development intervention itself impacted student achievement across the intervention years. The HLM results confirmed and explained the reasons for school-wide achievement and API gains for the five participating schools. Specifically, the results indicate that increases in student dose of the intervention (by proxy of their teachers' participation in the intervention) led to student gains in achievement, thus confirming that the intervention positively impacted student achievement, the primary goal of the intervention. Given that the intervention was designed to impact both students' achievement in math and their interest in and motivation for math by deliberately instructing and guiding teachers on how to use applied math in their classes as a means of increasing students' interest in, motivation for, and achievement in math, the intervention was highly successful for the participating teachers and for their students. The participating teachers also increased their math and science teaching efficacy over time, another research intervention goal.

The results of this research are highly generalizable to others who might attempt this structure and type of teacher intervention. First, it was built upon existing, impactful research as described in the literature review section of this manuscript. Second, both the teacher and student populations under study mirror that of many urban middle schools nationally (also described in the review of the literature). Third, given that the study results indicate that school-level effects (via the initial three-level HLM model), did not predict the achievement of the students in the study, the results suggest that the comprehensiveness of the intervention could apply to diverse school settings.

The results of this research are highly significant to educators, policy makers, and researchers. The findings provide ample evidence that teacher inservice professional development can positively impact students' mathematics achievement while simultaneously positively impacting teacher efficacy. These are significant results for teachers and particularly for their students. Policy makers and researchers in higher education have been highly skeptical about studying the impact of inservice teacher interventions on student achievement. This research negates such skepticism. Teacher educators will be highly interested in this research because (a) they often engage in inservice teacher professional development and will now have empirical rationale for continuing in such efforts, and

(b) these interventions could be utilized in preservice programs as a means of proactively supporting teachers and, ultimately, their students.

There are several factors that contributed to the intervention's success. The researchers achieved maximum fidelity of the intervention's implementation at the teacher level because of the comprehensive structure employed to guide and implement the teacher professional development. Additionally, using data-driven decision-making and using student achievement data as a metric for intervention content maximized the potential content impact for the students in the intervention schools. Finally, the fact that there was not teacher attrition in the research led to maximum doses of intervention by year three of the program with the eighth-grade students.

Study Limitations

The study is limited by several factors. The participating teachers were the primary administrators of the student outcome measures and therefore were responsible for collecting all student-level data. This was only modestly problematic because the research team optimized the structure for administration and data collection, made it similar to that of standardized K-12 assessment procedures, and kept theses structure constant semester-to-semester across the four study years. Additionally, the design of the intervention did not allow for pre-post comparison of teacher performance because lesson study was the means by which the TPOR scores were obtained, and the participating teachers only submitted one teaching video (used to assess performance using the TPOR) during the intervention time frame to study. This related to two factors: (a) the participating teachers needed to learn how to engage in lesson study as a critical structure for the intervention during which the videos were produced, and (b) the intervention period did not allow for substantial time to tape each teacher twice for comparison purposes.

Teacher attrition was not problematic in the intervention despite the socio-political climate of K-12 public schools associated with the nation's economic downturn during the intervention period. The participating district made a commitment to retain all of its math teachers during the intervention period because the need for stability of the teachers in the subject area was great considering the academic needs of the students and because the district recognized the importance of stability of teacher participants for the entire research period. The district's teachers' union and administration supported this decision at all levels.

COMPETING INTERESTS

There are no competing interests present for this manuscript.

FUNDING

This research was funded by the US Department of Education (CPEC ITQ 00347).

REFERENCES

Ben-Chaim, D. J., Fey, T., Fitzgerald, W. M., Benedetto, C., & Miller, J. (1998). Proportional reasoning among 7th-grade students with different curricular experiences. *Educational Studies in Mathematics*, *36*(3), 247–273. doi:10.1023/A:1003235712092

Borton, W. (1991, April 3–7). *Empowering teachers and students in a restructuring school: A teacher efficacy interaction model and the effect on reading outcomes* [Conference presentation]. Annual Meeting of the American Educational Research Association, Chicago, IL, United States.

Brown, A., Myers, J., & Collins, D. (2021). How pre-service teachers' sense of teaching efficacy and preparedness to teach impact performance during student teaching. *Educational Studies*, *47*(1), 38–58. doi:10.1080/03055698.2019.1651696

Bybee, R. W. (2010). Advancing STEM education: A 2020 vision. *Technology and Engineering Teacher*, *70*(1), 30–35.

Cai, J. (2003). What research tells us about teaching mathematics through problem solving? In F. Lester (Ed.), *Research and issues in teaching mathematics through problem solving*. National Council of Teachers of Mathematics.

Cai, J. (2010). Helping elementary students become successful mathematical problem solvers. In D. V. Lambdin & F. K. Lester (Eds.), *Teaching and learning mathematics: Translating research for elementary teachers* (pp. 9–14). National Council of Teachers of Mathematics.

California Department of Education. (2021). *Taking center stage—Act II: Ensuring success and closing the achievement gap for all of California's middle grades students*. Report to the Department of Education. https://www.cde.ca.gov/ci/gs/mg/tcsii-index.asp#:~:text=Taking%20Center%20Stage%E2%80%94Act%20II,Recommendations%20for%20Middle%20Grades%20Success

Caprara, G. V., Barbaranelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology*, *44*(6), 473–490. doi:10.1016/j.jsp.2006.09.001

Chiu, M. M., & Xihua, Z. (2008). Family and motivation effects on mathematics achievement: Analysis of students in 41 countries. *Learning and Instruction*, *18*(4), 321–336. doi:10.1016/j.learninstruc.2007.06.003

Cleary, T. J., & Chen, P. P. (2009). Self-regulation, motivation, and math achievement in middle school: Variations across grade level and math context. *Journal of School Psychology*, *47*(5), 291–314. doi:10.1016/j.jsp.2009.04.002 PMID:19712778

Cohen, J. (1992). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Coladarci, T. (1992). Teachers' sense of efficacy and commitment to teaching. *Journal of Experimental Education*, *60*(4), 323–337. doi:10.1080/00220973.1992.9943869

Darling-Hammond, L. (2012). *Powerful teacher education: Lessons from exemplary programs*. John Wiley & Sons.

Fuchs, L. S., Fuchs, D., & Bishop, N. (1992). Instructional adaptation for students at risk. *The Journal of Educational Research*, *86*(2), 70–84. doi:10.1080/00220671.1992.9941143

Fulton, K. P. (2012). 10 reasons to flip: A southern Minnesota school district flipped its math classrooms and raised achievement and student engagement. *Phi Delta Kappan*, *94*(2), 20–24. doi:10.1177/003172171209400205

Garet, M. S., Cronen, S., Eaton, M., Kurki, A., Ludwig, M., Jones, W., Uekawa, K., Falk, A., Bloom, H., Doolittle, F., Zhu, P., & Sztejnberg, L. (2008). *The impact of two professional development interventions on early reading instruction and achievement* [Report No. NCEE 2008-4030]. National Center for Education Evaluation and Regional Assistance, U.S. Department of Education. https://ies.ed.gov/ncee/pdf/20084030.pdf

Garet, M. S., Wayne, A. J., Stancavage, F., Taylor, J., Walters, K., Song, M., & Sepanik, S. (2010). *Middle school mathematics professional development impact study: Findings after the first year of implementation* [Report No. NCEE 2010-4009]. National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/pubs/20104009/pdf/20104010.pdf

Ghai, G., & Yaghi, H. (1997). Relationships among experience, teacher efficacy, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education*, *139*(4), 451–458. doi:10.1016/S0742-051X(96)00045-5

Glickman, C. D., & Tamashiro, R. T. (1982). A comparison of first-year, fifth-year, and former teachers on efficacy, ego-development, and problem-solving. *Psychology in the Schools*, *19*(4), 558–562. doi:10.1002/1520-6807(198210)19:4<558::AID-PITS2310190426>3.0.CO;2-F

Goldberg, P. D., & Bush, W. S. (2003). Using metacognitive skills to improve 3rd graders' math problem solving. *Focus on Learning Problems in Mathematics*, *25*(4), 36.

Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., Oliver, P. H., & Guerin, D. W. (2007). Multivariate latent change modeling of developmental decline in academic intrinsic math motivation and achievement: Childhood through adolescence. *International Journal of Behavioral Development*, *31*(4), 317–327. doi:10.1177/0165025407077752

Guthrie, J., & Klauda, S. (2014). Effects of classroom practices on reading comprehension, engagement, and motivations for adolescents. *Reading Research Quarterly*, *49*(4), 387–416. doi:10.1002/rrq.81 PMID:25506087

Hansen, M., & Gonzalez, T. (2014). Investigating the relationship between STEM learning principles and student achievement in math and science. *American Journal of Education*, *120*(2), 139–171. doi:10.1086/674376

Hanuscin, D., Donovan, D., Acevedo-Gutiérrez, A., Borda, E., DeBari, S., Melton, J., Le, T., Morrison, W., & Ronca, R. (2021). Supporting the professional development of science teacher educators through shadowing. *International Journal of Science and Mathematics Education*, *19*(S1, Suppl 1), 145–165. doi:10.1007/s10763-021-10154-5

Hashweh, M. Z. (2009). Teacher pedagogical constructions: A reconfiguration of pedagogical content knowledge. *Teachers and Teaching*, *11*(3), 273–292. doi:10.1080/13450600500105502

Helmbold, E., Venketsamy, R., & Heerden, J. V. (2021). Implementing lesson study as a professional development approach for early grade teachers: A South African case study. *Perspectives in Education*, *39*(2), 183–196. doi:10.18820/2519593X/pie.v39.i3.14

Hiebert, J., Gallimore, R., & Stigler, J. W. (2002). A knowledge base for the teaching profession: What would it look like, and how can we get one? *Educational Researcher*, *31*(5), 3–15. doi:10.3102/0013189X031005003

Hill, H. C., & Ball, D. L. (2004). Learning mathematics for teaching: Results from California's mathematics professional development institutes. *Journal for Research in Mathematics Education*, *35*(5), 330–351. doi:10.2307/30034819

Kayapinar, U. (2016). A study on reflection in in-service teacher development: Introducing reflective practitioner development model. *Educational Sciences: Theory & Practice*, *16*(5), 1671–1691. https://jestp.com/article-detail/?id=534

Klassen, R. M., Chong, W. H., Huan, V. S., Wong, I., Kates, A., & Hannok, W. (2008). Motivation beliefs of secondary school teachers in Canada and Singapore: A mixed methods study. *Teaching and Teacher Education*, *24*(7), 1919–1934. doi:10.1016/j.tate.2008.01.005

Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordon, K. A. (2011). Teacher efficacy research 1998–2009: Signs of progress or unfulfilled promise? *Educational Psychology Review*, *23*(1), 21–43. doi:10.1007/s10648-010-9141-8

Kohlmeier, J., Howell, J., Saye, J., McCormick, T., Shannon, D., Jones, C., & Brush, T. (2020). Investigating teacher adoption of authentic pedagogy through lesson study. *Theory and Research in Social Education*, *48*(4), 492–528. doi:10.1080/00933104.2020.1751761

Lazarides, R., Fauth, B., Gaspard, H., & Göllner, R. (2021). Teacher self-efficacy and enthusiasm: Relations to changes in student-perceived teaching quality at the beginning of secondary education. *Learning and Instruction*, *73*, 101435. Advance online publication. doi:10.1016/j.learninstruc.2020.101435

Lesh, R. A., & Doerr, H. M. (2013). *Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning, and teaching*. Routledge.

Lester, F. K. Jr. (2013). Thoughts About research on mathematical problem-solving instruction. *The Mathematics Enthusiast, 10*(1). *The Montana Math Enthusiast*, *12*(1-2), 245–278. doi:10.54870/1551-3440.1267

Lewis, C. C., Perry, R. R., & Hurd, J. (2009). Improving mathematics instruction through lesson study: A theoretical model and North American case. *Journal of Mathematics Teacher Education*, *12*(4), 285–304. doi:10.1007/s10857-009-9102-7

Lincoln, Y., & Guba, E. (1985). Naturalistic inquiry. *Sage (Atlanta, Ga.)*.

Listyani, E., Widjajanti, D. B., Susanti, M., Arliani, E., & Hidayati, K. (2008, July 31–August 1). *Development of mathematics high school teachers' competency through lesson study (a case study in Yogyakarta, Indonesia)* [Conference presentation]. International Conference on Lesson Study, Indonesia. https://staffnew.uny.ac.id/upload/131569343/penelitian/A.2+Endang+Listyani.pdf

Lotter, C., Yow, J. A., & Peters, T. T. (2014). Building a community of practice around inquiry instruction through a professional development program. *International Journal of Science and Mathematics Education*, *12*(1), 1–23. doi:10.1007/s10763-012-9391-7

Midgley, C., Feldlaufer, H., & Eccles, J. (1989). Change in teacher efficacy and student self- and task-related beliefs in mathematics during the transition to junior high school. *Journal of Educational Psychology*, *81*(2), 247–258. doi:10.1037/0022-0663.81.2.247

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical knowledge: A framework for teacher knowledge. *Teachers College Record*, *108*(6), 1017–1054. doi:10.1111/j.1467-9620.2006.00684.x

Mojavezi, A., & Tamiz, M. (2012). The impact of teacher self-efficacy on the students' motivation and achievement. *Theory and Practice in Language Studies*, *2*(3), 483–491. doi:10.4304/tpls.2.3.483-491

Moore, W., & Esselman, M. (1992, April 20–24). *Teacher efficacy, power, school climate and achievement: A desegregating district's experience* [Conference presentation]. Annual Meeting of the American Educational Research Association, San Francisco, CA, United States. https://files.eric.ed.gov/fulltext/ED350252.pdf

Morgan, N. W., & Morgan, L. M. (2013). An educational problem-solving strategy promoting team-work and interdisciplinary assessment through mathematical conceptual comprehension. *International Journal of Business and Social Science*, *4*(13), 11–15.

Mundry, S. (2005). Changing perspectives in professional development. *Science Education*, *14*(1), 9–15.

National Council of Teachers of Mathematics (NCTM). (2000). *Problem solving*. https://www.nctm.org/Research-and-Advocacy/research-brief-and-clips/Problem-Solving

Nurindah, A., Akil, M., & Jafar, B. (2019). Teachers' self-efficacy and performance in teaching literature in the interest-based classes at senior high school. *Journal of Language Teaching and Research*, *10*(6), 1271–1278. doi:10.17507/jltr.1006.16

OECD. (2023). *PISA 2022 assessment and analytical framework*. PISA, OECD Publishing. doi:10.1787/dfe0bf9c-en

Phelps, G., & Howell, H. (2016). Assessing mathematical knowledge for teaching: The role of teaching context. *The Montana Math Enthusiast*, *13*(1), 5. Advance online publication. doi:10.54870/1551-3440.1365

Protheroe, N. (2008). Teacher efficacy: What is it and does it matter? *Principal*, *87*(5), 42–45.

Ragusa, G. (2011). *Best practices in pedagogy and associated assessment* [Conference presentation]. National Science Foundation EEC Awardees.

Ragusa, G., Levonisova, S., & Huang, S. (2022). Improving middle school science achievement, literacy and motivation: A longitudinal study of a teacher professional development program. *Journal of STEM Education: Innovations and Research*, *23*(4), 22–35.

Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical linear models: Applications and data analysis methods. *Sage (Atlanta, Ga.)*.

Resnick, L. B. (1987). The 1987 presidential address: Learning in school and out. *Educational Researcher*, *16*(9), 13–20. https://www.jstor.org/stable/1175725

Riggs, I. M., & Enochs, L. G. (1990). Toward development of an elementary teachers' science teachers efficacy belief instrument. *Science Education*, *74*(6), 625–637. doi:10.1002/sce.3730740605

Roeser, R., Arbreton, A., & Anderman, E. (1993). *Teacher characteristics and their effects on student motivation across the school year* [Conference presentation]. American Educational Research Association.

Ross, J. A. (1998). The antecedents and consequences of teacher efficacy. In J. Brophy (Ed.), *Advances in research on teaching* (pp. 49–74). JAI Press.

Rozimella, Y. (2020). Developing teachers' professionalism through school-initiative lesson study. *European Journal of Educational Research*, *9*(4), 1513–1526. doi:10.12973/eu-jer.9.4.1513

Santagata, R., Zannoni, C., & Stigler, J. W. (2007). The role of lesson analysis in pre-service teacher education: An empirical investigation of teacher learning from a virtual video-based field experience. *Journal of Mathematics Teacher Education*, *10*(2), 123–140. doi:10.1007/s10857-007-9029-9

Schipper, T., Goei, S. L., de Vries, S., & van Veen, K. (2017). Professional growth in adaptive teaching competence as a result of lesson study. *Teaching and Teacher Education*, *68*, 289–303. doi:10.1016/j.tate.2017.09.015

Schmidt, W. H., Tatto, M. T., Bankov, K., Blömeke, S., Cedillo, T., & Cogan, L. (2007). *The preparation gap: Teacher education for middle school mathematics in six countries. Mathematics Teaching in 21st Century*. Center for Research in Mathematics and Science Education, Michigan State University.

Schoenfeld, A. (1994). Reflections on doing and teaching mathematics. In A. Schoenfeld (Ed.), *Mathematical thinking and problem solving* (pp. 53–69). Lawrence Erlbaum Associates.

Shulman, L. S. (2004). *The wisdom of practice: Essays on teaching, learning, and learning to teach*. Jossey-Bass.

Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: Effects of motivation, interest, and academic engagement. *The Journal of Educational Research*, *95*(6), 323–332. doi:10.1080/00220670209596607

Stigler, J. W., & Hiebert, J. (2009). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. Simon and Schuster.

Stigler, J. W., & Thompson, B. (2009). Thoughts on creating, accumulating, and utilizing shareable knowledge to improve teaching. *The Elementary School Journal*, *109*(5), 442–457. doi:10.1086/596995

Suryani, I., Maidiyah, E., Salasi, , & Mardhiah, M. Z. (2020). Students' mathematics problem-solving skills through the application of Problem-Based Learning model. *Journal of Physics: Conference Series*, *1460*(1), 012029. Advance online publication. doi:10.1088/1742-6596/1460/1/012029

Thanheiser, E., Browning, C., Moss, M., Watanabe, T., & Garza-Kling, G. (2010). Developing mathematical content knowledge for teaching elementary school mathematics. *Issues in the Undergraduate Mathematics Preparation of School Teachers, 1*. https://archives.pdx.edu/ds/psu/11157

Tschannen-Moran, M., & Barr, M. (2010). Fostering student learning: The relationship of collective teacher efficacy and student achievement. *Leadership and Policy in Schools*, *3*(3), 189–209. doi:10.1080/15700760490503706

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*(7), 783–805. doi:10.1016/S0742-051X(01)00036-1

Uygun, N., & Tertemiz, N. I. (2014). Effects of problem-based learning on student attitudes, achievement and retention of learning in math course. *Education in Science*, *39*(174), 75–90. doi:10.15390/EB.2014.1975

Wei, R. C., Darling-Hammond, L., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the United States and abroad*. National Staff Development Council.

Wenglinsky, H. (2000). *How teaching matters: Bringing the classroom back into discussions of teacher quality*. A Policy Information Center Report. The Milken Family Foundation and Educational Testing Service. https://files.eric.ed.gov/fulltext/ED447128.pdf

Wilson, M. (2011). *Constructing measures: An item response modeling approach*. Lawrence Erlbaum Associates Publishers.

Windschitl, M. (2009, February 5–6). *Cultivating 21st century skills in science learners: How systems of teacher preparation and professional development will have to evolve* [Conference presentation]. National Academies of Science, Workshop on 21st Century Skills, Washington, DC, United States.

Xiao, F., & Sun, L. (2021). Students' motivation and affection profiles and their relation to mathematics achievement, persistence, and behaviors. *Frontiers in Psychology, 11*, Article 533593.

APPENDIX A

Table 7. Variables and Descriptors in HLM Models

Variables and Acronyms in Student-Level Models	
MATH BENCHMARK	An average score of four quarters of math benchmark tests in one academic year
MATH SST	A State Standards Test score for Middle School Math in each grade
MATH INTRINSIC MOTIVATION	Student intrinsic motivation for learning Math
MATH EXTRINSIC MOTIVATION	Student extrinsic motivation for learning Math
MATH MOTIVATION	Student overall motivation for learning Math
STUDENT DOSE	Amount of interventions received by students
GENDER	Male or Female
HISPANIC	Race/Ethnicity
AFRICAN AMERICAN	Race/Ethnicity
WHITE	Race/Ethnicity
FILIPINO	Race/Ethnicity
Variables and Acronyms in Teacher-Level Models	
TEACHER DOSE	Amount of training a teacher received from the professional development program
TEACHER PERFORMANCE	Teacher performance score in math/science
TEACHER EFFICACY	Teachers' efficacy to teach math/science
SINGLE-SUBJECT CREDENTIAL	Single-subject (versus multiple-subject credential; dummy coded)
TEACHER EXPERIENCE	Number of years of experience in teaching math/science
MATH INTRINSIC MOTIVATION	Centered math intrinsic motivation score using the group mean
MATH EXTRINSIC MOTIVATION	Centered math extrinsic motivation score using the group mean
MATH MOTIVATION	Centered overall math motivation score using the group mean

APPENDIX B

General Abbreviations

API: Academic Performance Index
HLM: Hierarchical linear modeling
MPPTP: Measuring perception of pre-service teachers' preparedness
MPS: Mathematical Problem Solving
MTEBI: Math Teaching Efficacy Beliefs Instrument
NCTM: National Council of Teachers of Mathematics
PD: Professional Development
PI: Program Improvement
PISA: Program for International Student Assessment
RTOP: Reformed Teacher Observation Protocol
SST: State Standards Tests
STEM: Science, technology, engineering, and mathematics
TLT: Teacher Leadership Team
TPOR: Teacher Performance Observational Rubric
TSES: Teacher sense of efficacy scale

Gisele Ragusa received her Ph.D. from the University of Southern California in Los Angeles, California. She is currently a Professor of Engineering Education at the University of Southern California. Before becoming a professor, she was a K-12 teacher for 14+ years. Her research interests include science, technology, engineering and mathematics (STEM) education, engineering innovation, college access, STEM PK-12 education and teacher education, student mentorship, STEM literacy education, human-centered design, socially assistive robotics, as well as assessment and measurement. Dr. Ragusa received the U.S. Presidential Award for Excellence in Science, Mathematics, and Engineering Mentoring (PAESMEM) in 2019.

Shaobo Huang is a professional engineer in Canada. She received a Ph.D. degree in Engineering Education from Utah State University and collaborated with other authors on this paper as a postdoctoral associate researcher at the University of Southern California. She has over eight years of teaching and/or research experience in engineering education. She is currently an Assistant Professor in the Ron and Jane Graham School of Professional Development with a joint appointment in the Department of Mechanical Engineering at the University of Saskatchewan in Canada. Dr. Huang is one of the key members on the RE-ENGINEERED First-Year Program at the University of Saskatchewan.

Svetlana V. Levonisova earned her Ed.D. degree in Administration in Higher Education from the University of Southern California. She participated in this research paper as a postdoctoral scholar. She is an international educator with 20 years of experience in higher education including teaching and research that she started in Russia. Currently, she is an Assessment Coordinator at the Attallah College of Educational Studies at Chapman University. She is working for the Office of Accreditation and Assessment to coordinate and support state and national accreditation processes for seven graduate and one undergraduate programs.