EXAMINING THE IMPACT OF SCHOOL CULTURE, TEACHER PROFESSIONAL LEARNING, AND INSTRUCTIONAL PRACTICES ON MIDDLE GRADES STUDENTS' STEM OUTCOMES

A Dissertation

by

KIMBERLY BODDIE WRIGHT

Submitted to the Office of Graduate and Professional Studies of

Texas A&M University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Hersh C. Waxman
Comittee Members,	Cheryl J. Craig Michael A. de Miranda Timothy P. Scott
Head of Department,	Michael A. de Miranda

December 2018

Major Subject: Curriculum and Instruction

ABSTRACT

This dissertation examined effective STEM instructional strategies as well as the ways in which the four aspects of middle grades students' STEM pipeline experiences are impacted by malleable school and classroom factors. Specifically, the dissertation included three studies that focused on determining: (a) STEM instructional interventions shown to be effective for middle grades students, (b) the impact of school factors on middle grades STEM teachers' use of effective STEM practices, and (c) the impact of teachers' perceptions on their use of effective STEM practices. Study predictors were mainly derived from the *Contexts for Teachers' Learning* framework. Study one found that, on average, students involved in STEM interventions performed 0.424 standard deviations higher than students in the control group in experimental studies or prior to the intervention in pre/post studies. Study two found that building teachers' professional capacity, as well as providing coherent instructional guidance, leadership opportunities, and adequate time and funding had statistically significant and positive impacts on teaching practice. Study three found that building teachers' professional capacity and providing adequate time and/or funding resulted in statistically significant positive impacts on instructional practices. Finally, qualitative analysis of teachers' responses in study three highlighted the importance of the availability of instructional technology, the importance of developing teachers' professional capacity, and a potential need to differentiate professional learning efforts by years of teaching experience and STEM subject area.

DEDICATION

I dedicate this dissertation to family in its many forms. Mom, Dad, Bryce, and Papa– none of this happens without the grace and love that started with each of you. Mom, there aren't enough words to tell you how much your support means to me and my tribe of crazies. This dissertation is possible because of your love, sacrifice, and hard work. Bryce, it's your turn now. You've got this! To my extended family in the Wrights, Buchanans, and Betheas– your support means the world. Old Army! To all of my K-12 school families at Dellview, Olmos, Kemp, Davila, and Wellborn– you taught, and continue to, teach me what it means to be a teacher. I promise to never stop becoming and fighting on your behalf. To my Education Research Center family– thank you for seeing the researcher in me and giving me time, patience, and mentorship to develop a lens from which to view education in a different way. To Dr. Jacqueline Stillisano– if not for the opportunities you have extended to me over the last decade, I would not be where I am today. Thank you for your constant belief, encouragement, and mentorship. To my TLAC graduate student family, and most especially to Mario Itzel Suárez– thank you for walking this PhD road with me. It has been my greatest honor to learn and grow alongside you these last three years. I cannot wait to see where this road leads us. To my K-12 public school teacher family in Texas and across the country– I hear you, I see you, and your voice matters. Never lose faith that your work, our work, is among the most important work. Finally, I dedicate this dissertation to my 1817 Bee Creek Drive family. Terrell, Ella, and Webb, my heart begins and ends with the three of you. I hope this final PhD work makes you proud and happy that we finally have all of our nights and weekends back. Thank you for the sacrifice of time that you three made to help me get to this finish line.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Hersh Waxman, and my committee members, Dr. Michael DeMiranda, Dr. Cheryl Craig, and Dr. Timothy Scott, for their guidance and support throughout the course of this research.

Thanks also go to my friends and colleagues in the Education Research Center for making me an education researcher.

Thank you to the faculty and staff of the College of Education and Human Development's Department of Teaching, Learning and Culture for making my time at Texas A&M University a great experience.

A final thanks goes to YouTube®, iMessage®, Bitmoji®-enabled sarcasm, coffee in all of its glorious forms, and Chuck Huber. This dissertation is largely a product of these crucial ingredients.

CONTRIBUTORS AND FUNDING SOURCES

This work was supported by a dissertation committee consisting of Professor Hersh Waxman and Drs. Michael DeMiranda and Cheryl Craig of the Department of Teaching, Learning, and Culture and Professor Timothy Scott of the Department of Biology.

The coding of studies analyzed for Chapter III was conducted in part by Manjari Banerjee and Jeffrey Keese of the Department of Teaching, Learning, and Culture.

All other work conducted for the dissertation was completed by the student independently.

There are no outside funding contributions to acknowledge related to the research and compilation of this document.

NOMENCLATURE

STEM	Science, Technology, Engineering, and Mathematics
NCLB	No Child Left Behind Act of 2001
ESSA	Every Child Succeeds Act of 2015
TPL	Teacher Professional Learning
PD	Professional Development

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

Over 30 years ago, *A Nation at Risk* (National Commission on Excellence in Education, 1983, p. 5) lamented the "rising tide of mediocrity" in U. S. schools, particularly in mathematics and science. In the decades since, determining how to develop and maintain a high-quality science, technology, engineering, and mathematics (STEM) workforce continues to be an issue of paramount importance. In addition to its economic importance, the promotion of STEM literacy is considered not just necessary for continued economic success, but "a democratic ideal worthy of focused attention, significant resources, and continuing effort" (National Research Council, 2012, p. 277). With over 80% of the 30 fastest growing occupations in the United States in 2016 in STEM-related fields (U. S. Bureau of Labor Statistics, Employment Projections Program, 2017), continued K-12 public school improvement in STEM is viewed as fundamental to increasing the international competitiveness of U. S. graduates (National Academy of Sciences, National Academy of Engineering, and Institute of Medicine, 2007).

In a landmark report a decade ago, members of the National Academies of Sciences, National Academy of Engineering, and Institute of Medicine (2007) put forth a joint report detailing 10 recommendations targeted at improving science and technology in the U. S. in order for the nation to compete in the global community of the 21st century. Chief among the recommendations was improving K-12 mathematics and science education in order to graduate more high school students capable of obtaining undergraduate degrees in STEM. Despite progress around the 2007 recommendations, a great deal of concern still surrounds the state of STEM education in the U. S., as well as the preparation and instructional practices of STEM teachers (Carnegie Commission for Mathematics and Science Teaching, 2009; Coble, 2012; National Academy of Sciences, National Academy of Engineering, & Institute of Medicine, 2010; Presley & Coble, 2012; Wilson, 2013). With less than half of women and men across the nation persisting through STEM degrees, such as engineering (Hill, Corbett, & St. Rose, 2010), and less than 50% of the degrees conferred by U. S. postsecondary institutions occurring in STEM fields (McFarland et al., 2017), examining factors that impact students' persistence to and through STEM undergraduate majors continues to be an issue of utmost concern.

In addition to an overall shortage of students graduating with a STEM degree, there remain substantial gaps in the representation of different ethnic groups among STEM graduates. Despite progress along gender lines for female graduates in STEM, with females representing 63% of the 2014-15 STEM graduates from U. S. postsecondary institutions (McFarland et al., 2017), gaps remain between percentages of White and underrepresented minority STEM graduates. Underrepresented minority students, specifically Black/African American, Hispanic/Latinx, and American Indian/Alaska Native students, continue to be disproportionately represented in STEM fields (National Action Council for Minorities in Engineering, 2013). For example, in 2014, White students made up over 60% of STEM graduates, while Black and Hispanic student groups made up eight percent each of the STEM graduates. The most recent numbers of Black and Hispanic STEM graduates closely mirror figures from prior research reporting the underrepresentation of Blacks and Hispanics in the STEM workforce. A 2011 report from the U.S. Department of Commerce's Economics and Statistics Administration revealed that though Blacks and Hispanics accounted for 11% and 14% of the overall workforce in 2009, each of the groups only accounted for six percent of STEM workers (Beede, Julian, Khan, Lehrman, McKittrick, Langdon, & Doms, 2011).

A large body of research has established clear linkages between fixed student, teacher, and school factors; such as student race/ethnicity or SES (Berryman, 1983; Hanson, 1996; Hinojosa, Rapaport, Jaciw, LiCalsi, & Zacamy, 2016; Oakes, 1990), teacher preparation (DeAngelis & Presley, 2011; Sass, 2015; Wilson, 2011), school demographics (Aschbacher, Li, & Roth, 2010; Change the Equation, 2016; National Research Council, 2012), and student performance in STEM disciplines. The bleak conclusions drawn from much of this research have been regularly emphasized in reports and standards documents calling for changes to K-12 STEM education (National Academy of Sciences, National Academy of Engineering, and Institute of Medicine; 2007; National Academy of Sciences, National Academy of Engineering, and Institute of Medicine; 2010; National Action Council for Minorities in Engineering, 2013). With titles that have historically emphasized impending doom should educators fail to heed the "risk" or "storm" associated with each subsequent report or set of standards, reports such as the National Academies' *Rising Above the Gathering Storm: Energizing and*

Employing America for a Brighter Economic Future (2007) and its follow-up report, *Rising Above the Gathering Storm: Revisited, Rapidly Approaching Category 5* (2010) provide a myriad of recommendations for what K-12 STEM teachers should know and be able to do, as well as how schools should structure learning for both STEM teachers and their students.

As a result of such sustained external focus and scrutiny, K-12 STEM education, when viewed as a collection of subjects, is guided by close to 10 sets of national standards, written both for students and teachers. Just over 30% of U. S. states have adopted the *Next Generation Science Standards* (NGSS; NGSS Lead States, 2013) for science and engineering, while 84% of U. S. states follow the *Common Core State Standards for Mathematics* (CCSSM; National Governors Association, the Council of Chief State School Officers [CCSSO], & Achieve, Inc., 2008). In addition to student standards, most states have separate sets of teacher proficiencies governing certification in each STEM subject, along with additional sets of standards for teacher evaluation; while organizations, such as the National Board for Professional Teaching Standards (2016) and the CCSSO's Interstate Teacher Assessment and Support Consortium (2011) provide additional sets of teaching standards at the national level. Finally, several STEM-related professional organizations, such as the National Academies and the American Statistical Association, have also independently created lists of standards for STEM teachers and students (American Statistical Association, 2007; National Academy of Sciences, National Academy of Engineering, and Institute of Medicine, 2007;

National Academy of Sciences, National Academy of Engineering, and Institute of Medicine, 2010).

Though recommendations for STEM education abound, there remains a lack of strong evidence as to which of the myriad recommendations are most effective in specific K-12 contexts and varied aspects of students' STEM-related experiences, as well as the degree to which recommendations are regularly implemented in STEM classrooms across the U. S. An overabundance of standards combined with a lack of evidence supporting their effectiveness has led to an increased focus on determining what core factors at the school and classroom level can be adjusted to contribute to more students, specifically more underrepresented minority students, entering and remaining in the STEM pipeline (Aschbacher et al., 2014; Brotman & Moore, 2008; Institute of Education Sciences, U. S. Department of Education & the National Science Foundation, 2013; Maltese & Tai, 2011; National Research Council, 2012; Wilson, 2011).

Research on enhancing students' STEM experiences in order to increase the numbers of underrepresented students and women entering the STEM pipeline has been an area of intense research for several decades. The STEM pipeline metaphor, first introduced by Berryman (1983) in an investigation of the underrepresentation of minorities and women in post-secondary degrees in STEM fields, suggested that the path to STEM was an ever-narrowing conduit through which a talented pool of students must pass to enter a STEM major or career. The majority of the factors studied were fixed factors, including years living in the U. S., parental educational attainment, number of math courses taken, and membership in a particular racial or ethnic subgroup. Berryman concluded that two main factors, persistence through the pipeline and choice of field, influenced students' matriculation to STEM majors.

In a follow-up study of the trends in the participation of female and minority representation in STEM fields, Oakes (1990) extended the STEM pipeline factors to include not just achievement and interest, but opportunities to study STEM. Like Berryman, Oakes' work examined research focused on how female and minority participation in STEM is influenced by fixed factors, such as parental education, socio-economic status, and parent participation in school. The study's main conclusion was that limited opportunities to participate in activities and experiences that generate interest and achievement in STEM inhibit greater matriculation through the STEM pipeline. Hanson (1996) combined the pipeline frameworks of Berryman and Oakes to include four aspects of students' experiences in the STEM pipeline: achievement, access, attitude, and activities. In the two decades since, these four aspects of students' experiences have been studied extensively, with clear themes emerging from much of the research regarding how and at what point in a students' K-12 trajectory the experiences may be influenced.

Literature Review

STEM Achievement

Achievement in K-12 STEM subjects, more than any aspect of STEM pipeline experiences, has long been considered a strong predictor of student matriculation through the STEM pipeline to an undergraduate STEM major. In a survey with a national sample of U. S. undergraduate students enrolled in introductory courses (n = 6,882), students who reported higher middle school mathematics grades had 1.5 times higher odds per letter grade to report a STEM career interest at the university level (Dabney, Tai, Almarode, Miller-Friedmann, Sonnert, Sadler & Hazari, 2012). A longitudinal study with 4,700 students in U.S. schools who participated in National Education Longitudinal Study of 1988 that followed students for twelve years from eighth grade through postsecondary found that students who earned higher scores on eighth-grade science and mathematics achievements were more likely to complete degrees in STEM (Maltese & Tai, 2011).

Student STEM achievement has been found to not only impact students' choice of STEM majors and careers, but also has been found to have a positive association with other aspects of students' experiences in the STEM pipeline, namely students' attitude towards STEM. Using data from the 2003 Trends in International Mathematics and Science Study (TIMSS), researchers found that eighth-grade students' science test scores had a positive and significant association with science career aspirations (Riegle-Crumb, Moore, & Ramos-Wada, 2010). Large achievement gaps were also found between white males and Black and Hispanic males in the TIMSS sample, consistent with racial/ethnic gaps in STEM undergraduate degrees (McFarland et al., 2017).

Attitude Towards STEM

Along with achievement, students' attitudes towards STEM subjects and careers is an aspect of STEM pipeline factors that has received a lot of attention in research. Berryman (1983) and Oakes (1990) first identified the influence of students not only being capable in STEM subjects, but also making the choice to engage in further study.

Most studies of students' STEM attitudes are strongly influenced both by social-cognitive theory (Bandura, 2001) and expectancy-value theory (Eccles & Wigfield, 2002), focusing on interactions between students' interests and motivation in STEM subjects and the intermediate and/or long-term impacts on students' choice, or intended choice, of STEM careers. A longitudinal study that tracked middle school students (n = 3,359) from eighth grade to age 30 found that students who were interested in a science career in middle school were between 1.9 and 3.4 times more likely to have a science-related career (Dabney et al., 2012). A study of the science self-perceptions of another group of eighth-grade students (n = 493) at a diverse urban middle school in California found that students' self-perceptions of their ability to do science and their perceived value of science predicted career interest in science (Aschbacher et al., 2014).

Similar to the influence of students' attitudes on science career interests and achievement, studies have also shown students' perceptions of mathematics to be highly influential in persistence and goal-orientation in the subject. In a study with students in grades seven through 11 (n = 759), students' effort in mathematics was mainly explained by their beliefs in their mathematics competency and mastery-oriented goals (Chouinard, Karsenti, & Roy; 2007). Mathematics mastery goal-orientation in students, contrasted with performance-approach and work avoidance orientations, are also associated with lower levels of anxiety and more use of help-seeking behavior in students (Federici, Skaalvik & Tangen, 2015). Similar to achievement gaps in STEM, both gender and racial/ethnic gaps have been found in the STEM-related attitudes of females and minority students, with the STEM attitudes of underrepresented female students, more than other groups, particularly vulnerable to decline as students transition from middle to high school (Ing & Nylund-Gibson, 2017).

Students' Access in STEM

Student access in STEM is a broad idea encompassing students' access not just to material resources, but also adult guidance, content, instruction, and teacher expectations (Oakes, 1990). A recent report from the results of the 2014 Technology and Engineering Literacy Survey (Change the Equation, 2016) highlighted the urgent need for schools to increase students' access to facilities and materials that provide students with opportunities to build things and take things apart. Of the 21,500 eighth-grade students surveyed across the U. S., less than 20% of students surveyed had access to materials and/or experiences that allowed them to engage in engineering practices such as building and testing models, taking things apart, or using a variety of tools to determine which is superior for a particular task.

Of the four STEM pipeline factors, access may be the one that is most variable across different racial and ethnic groups. Students from underrepresented minority groups are less likely than white students to have access to things ranging from advanced courses (Aschbacher et al., 2010; Hill et al., 2010; National Research Council, 2012) to STEM-related career guidance and technology (Change the Equation, 2016; Hinojosa et al., 2016). From discourse patterns in classrooms to access to advanced course-taking and advising, student access to STEM experiences has been highlighted as one of the four most important areas in the expansion of the STEM pipeline for underrepresented minority students (Committee on Underrepresented Groups and the Expansion of the

Science and Engineering Workforce Pipeline; Committee on Science, Engineering, and Public Policy; Policy and Global Affairs; National Academy of Sciences; National Academy of Engineering; Institute of Medicine; 2011).

Students' STEM-related Activities

Research on STEM-related activities centers mainly around student participation in STEM extracurricular experiences and types of STEM classroom instruction. Extracurricular science experiences have been found to positively predict both science attitudes and interest (Brotman & Moore, 2008; Dabney et al., 2012). A 2012 survey of a nationally representative sample of undergraduate students (n = 6,882) found that students who participated in out-of-school time (OST) science activities at least a few times per year were 1.5 times more likely to choose a STEM major than students who did not participate or who participated in OST activities less frequently (Dabney et al., 2012). A retrospective study of 33 ethnically diverse high school students who were very interested in STEM as tenth-grade students found that the majority of high-achieving persisters, or those students who were both high achievers and still interested in pursuing a STEM undergraduate major as high school seniors, participated in hands-on extracurricular experiences in places such as labs and hospitals where they had opportunities to engage in real world STEM experiences and interact with doctors and scientists (Aschbacher et al., 2010). However, the study authors noted that, despite their positive association with students' plans to major in STEM, opportunities to participate in extracurricular STEM activities were also influenced by external factors, such as socio-economic status (SES) and family support. For example, low- to mid-SES students were more likely to have jobs after school that prohibited them from participating in extracurricular activities or parents who were less aware of opportunities for extracurricular STEM experiences and/or unable to pay related expenses.

In-school activities including specific instructional practices and learning formats, as well as technology integration, have been the focus of a number of recent studies. Instructional practices linked to increasing students' interest and achievement in STEM have been widely studied and include things such as hands-on experiences, open-ended tasks, relevant contexts, cooperative learning (Brotman & Moore, 2008; Christensen, Knezek, & Tyler-Wood, 2015; Dare & Roehrig, 2016; Nugent, Barker, Welch, Grandgenett, Wu & Nelson, 2015; Riegle-Crumb, King, Grodsky, & Muller, 2011). In addition, technology activities, such as virtual group experiences (Brown, Concannon, Marx, Donaldson, & Black, 2016), as well as the integration of technology with other activities, such as project-based learning (Hansen & Gonzalez, 2014; Kim, 2016), result in increases in students STEM attitudes, including self-perceptions, efficacy, and persistence.

There is emerging evidence that effective STEM instructional activities have an even greater influence on female and minority students (Colvin, Lyden, Leon de la Barra, 2013; Dare & Roehrig, 2016). For example, a study of fifth- and sixth-grade female students (n = 45) emphasizing collaborative approaches to civil engineering projects resulted in increased student views of females as engineers (Colvin et al., 2013). Another study concluded that female students who participated in class discussions, hands-on activities, and experiences were more likely to perceive that physics is connected to everyday life (Dare & Roehrig, 2016). In contrast to positive findings related to hands-on, collaborative STEM activities, another study found that students who were considered "lost potentials" (Aschbacher et al., p. 569), or students who showed early initial interest in STEM but no longer wished to pursue a STEM major by twelfth grade, reported few hands-on activities or meaningful projects. The majority of students in the "lost potentials" group were black or Hispanic, suggesting a possible connection between race/ethnicity and quality STEM-related activities.

Context of the Present Study

This dissertation examines the ways in which the four aspects of students' STEM pipeline experience (achievement, attitudes, access, and activities) are impacted by malleable school and classroom factors. The present study focuses on middle grades students and teachers due to overwhelming evidence in the extant literature that the middle grades is a time when students' attitudes regarding STEM fields and careers are most subject to change (Aschbacher et al., 2014; Catsambis, 1995; Christensen & Knezek, 2016; Hinojosa et al., 2016; Ing & Nylund-Gibson, 2017; Nugent et al., 2015; Oakes, 1990). Though there is some inconsistency in how middle grades is defined, the present study utilizes the U. S. Department of Education's Institute for Education Sciences operationalization of middle grades as grade levels five through eight (Snyder, de Brey, & Dillow, 2016).

A great deal of research has been conducted on one or more aspects of students' middle grades STEM pipeline experiences. However, the majority of studies rely on student-reported data. Few studies provide an account of the degree to which teachers perceive that their schools and classrooms are providing students with the types of experiences found in research to be influential in their matriculation to STEM undergraduate majors and careers, as well as how teacher perceptions of their practices contribute to student experiences and outcomes in STEM.

In addition, much of the current research focuses on factors fixed factors that cannot be manipulated at the school level to directly impact students. For example, Hanson's (1996) consideration of external factors acting on the STEM pipeline was limited mostly to fixed ideas such as family structure, school characteristics, and courses taken. Though Hansen's framework did include an examination of a few malleable factors, such as teacher and student attitudes, missing from the framework were mechanisms that research has more recently identified as impactful in general school improvement efforts (Allensworth, Ponisciak, & Mazzeo, 2009; Brophy, Klein, Portsmore, & Rogers, 2008; Bryk & Schneider, 2002; DeAngelis & Presley, 2011; Maltese & Tai, 2011), such as school culture (i.e., level of collaboration, classroom autonomy), professional learning (i.e., amount and types of professional development in STEM), and specific instructional practices (i.e., classroom discourse, hands-on learning experiences, real world connections).

In 2013, the Institute for Education Sciences, U. S. Department of Education, and the National Science Foundation emphasized the importance of both alterable and ground-level factors with the establishment of a focus on malleable school and classroom factors as a requirement of all early stage or exploratory research programs seeking federal funding from either organization. The institutions defined malleable factors as, "…factors that are alterable, such as children's behaviors; technologies; education programs; policies; and practices," (2013, p. 12). A recent review of factors impacting Hispanic student success in STEM (Hinojosa et al., 2016) also emphasized the importance of focusing on indicators that are: (a) predictive of student success in STEM, (b) malleable, and (c) actionable at the school or district, rather than state or federal, levels. However, less than a quarter of the studies reviewed focused on school or teacher characteristics and only one study focused on STEM pedagogy. None of the studies reviewed focused on differences in school- or classroom-level predictors between Hispanic and non-Hispanic students. In order for schools and classroom teachers to maximize the effectiveness of ground-level efforts to broaden the STEM pipeline, it is critical to investigate the impact of school culture and teacher professional learning on teachers' use of effective STEM practices.

The purpose of this multiple-article dissertation is to examine: (a) instructional practices that positively impact middle grades students' STEM-related achievement, activities, access, and attitudes, (b) the impact of school culture on middle grades STEM teachers' use of effective STEM practices, and (c) the impact of teachers' professional learning on their use of effective STEM practices. The study uses a multi-tiered approach to examine effective middle grades STEM practices, as well as the extent to which school culture and teacher professional learning impact the use of effective STEM practices in schools nationwide, as well as within the state of Texas. The first study uses meta-analysis to examine which aspects of middle grades instructional practices are identified in the research literature as most effective in the development of the two of the four domains identified in Hanson's (1996) interpretation of the STEM pipeline: achievement and attitude. The second study utilizes secondary data analysis to explore teacher and school factors that explain variation in teacher self-reports of effective STEM practices in a nationally representative sample of middle grades STEM teachers. Finally, the third study investigates teacher and school factors that explain variation in teacher self-reports of effective STEM practices in middle grades schools in the state of Texas. Figure 1.1 provides an overview of the conceptual framework of the dissertation.

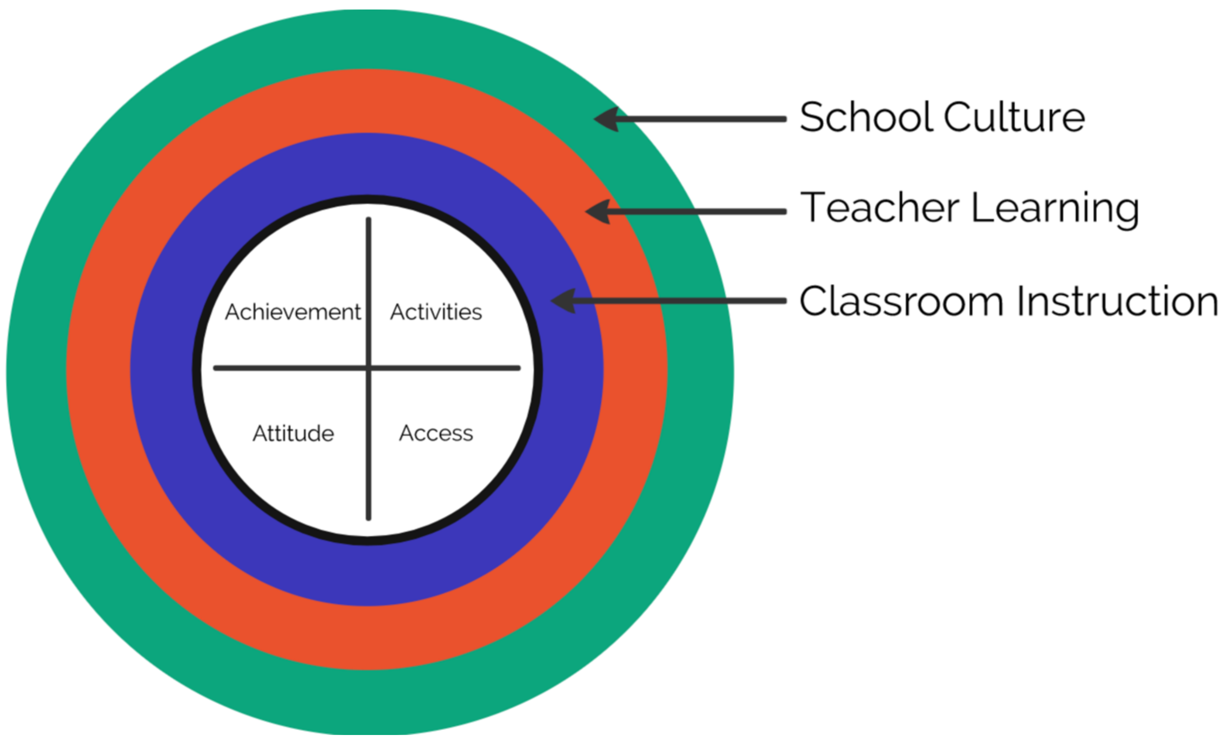


Figure 1.1 *The Impact of School and Classroom Factors on Middle Grades Students' STEM Pipeline Experiences.*

The study's findings provide information for policy makers, district and school-level administrators, and K-12 classroom teachers regarding which malleable school and teacher factors are most impactful in increasing the numbers and types of students matriculating into STEM fields. The findings also highlight the degree to which effective STEM practices are present in middle grades classrooms across the U. S. and Texas, as well as the degree to which practices differ in high-performing, high poverty schools and low-performing, high poverty schools. Finally, the study adds to the research base on how STEM practices identified in the research literature are utilized in middle grades classrooms.

Method

This study utilizes a multiple journal article format. Each manuscript is provided in its entirety, including a title and overview of research questions, data sources and instruments, data analyses, and conclusion. Texas A&M University Institutional Review Board approval has been obtained for all proposed studies (TAMU IRB Number: IRB2017-0770D).

CHAPTER II

A META-ANALYSIS OF FACTORS IMPACTING MIDDLE GRADES STUDENTS' ACHIEVEMENT AND ATTITUDES IN STEM

As our nation and world continues its shift to a knowledge-based economy in which individuals must be equipped with the ability to gather and analyze information from a variety of media to solve multi-faceted problems, the promotion of a populace literate in science, technology, engineering, and mathematics (STEM) continues to be a topic of great national interest. The important skills inherent in the STEM disciplines extend beyond content knowledge and include process-based thinking skills such as problem ideation and problem solving, persistence, and creativity (Bailey, Kaufman, & Subotic, 2015). In addition to the fact that over 80% of the 30 fastest growing occupations in the United States in 2016 were in STEM-related fields (U. S. Bureau of Labor Statistics, Employment Projections Program, 2017), the types of analytical skills present in STEM disciplines are in demand across a diverse array of jobs, including construction, manufacturing, public administration, and management (Rothwell, 2013).

The continued growth of a STEM-focused economy has resulted in an intense focus on the development of K-12 students' STEM competencies in order to increase numbers of STEM-skilled students matriculating into higher education or directly into the STEM workforce (Carnegie Commission for Mathematics and Science Teaching, 2009; Coble, 2012; National Academy of Sciences, National Academy of Engineering, and Institute of Medicine, 2007; National Research Council, 2012; NGSS Lead States, 2013; Wilson, 2013). As educators seek to determine important leverage points for the development of STEM competencies along the K-12 trajectory, the middle grades, generally agreed upon as grades five through eight (U. S. Department of Education, Institute for Education Sciences, National Center for Education Statistics, 2015), has emerged as a time when students' attitudes regarding STEM fields and careers are most subject to change (Aschbacher, Ing, & Tsai, 2014; Catsambis, 1995; Christensen & Knezek, 2016; Hinojosa, Rapaport, Jaciw, LiCalsi, & Zacamy, 2016; Ing & Nylund-Gibson, 2017; Nugent, Barker, Welch, Grandgenett, Wu, & Nelson, 2015; Oakes, 1990). Middle grades students' achievement in, and attitude toward, STEM disciplines has been shown to predict later achievement and matriculation to STEM undergraduate majors and STEM careers (Aschbacher et al., 2014; Dabney, Tai, Almarode, Miller-Friedmann, Sonnert, Sadler & Hazari, 2012; Maltese & Tai, 2011).

Though recommendations for STEM education abound, there remains a lack of strong evidence as to which of the myriad recommendations are most effective in middle grades contexts. Several recent meta-analytic studies of STEM interventions and programs have focused on different types of STEM programs, teaching strategies, or subject areas with positive findings across grades K through 12, with effect sizes ranging from small to large effect sizes across studies (An, 2013; Cheung & Slavin, 2013; Schroeder, Scott, Tolson, Huang, & Lee, 2007). Though these studies provide evidence as to the effectiveness of STEM education across K-12, including the middle grades, the differential operationalization of grade bands across the studies makes it difficult to determine specifically how the included studies impacted middle grades students. The

An study (2013) found STEM programs to have a large effect (0.880) on middle school students, however, the author did not include information on how middle school was defined in terms of included grade levels. In a meta-analysis of the impact of educational technology applications on mathematics achievement, Cheung and Slavin (2013) found a small effect size (0.14) for secondary students, operationalized as students in grades six through 12. Finally, in a meta-analysis of the impact of science teaching strategies on science achievement, Schroeder et al. (2007) found instruction to have a moderate (0.66) impact on elementary students' science achievement, with elementary including students in grades K through 8.

In order to provide a comprehensive look at the impact of STEM interventions in the middle grades, the present meta-analysis includes only studies conducted with students in grades five through eight, thus providing information on the degree to which various interventions are impactful for middle grades students specifically. In addition, the study includes multiple STEM subject areas and outcomes (i.e., achievement and attitudes).

Theoretical Framework

Research on enhancing students' STEM experiences in order to increase the numbers of underrepresented students and women entering the STEM pipeline has been an area of intense research for several decades. The STEM pipeline metaphor, first introduced by Berryman (1983) in an investigation of the underrepresentation of minorities and women in post-secondary degrees in STEM fields, suggested that the path to STEM was an ever-narrowing conduit through which a talented pool of students must pass to enter a STEM major or career. Berryman concluded that two main factors, persistence through the pipeline and choice of field, influenced students' matriculation to STEM majors.

In a follow-up study of the trends in the participation of female and minority representation in STEM fields, Oakes (1990) extended the STEM pipeline factors to include not just achievement and interest, but opportunities to study STEM. Oakes' main conclusion was that limited opportunities to participate in activities and experiences that generate interest and achievement in STEM inhibit greater matriculation through the STEM pipeline. Hanson (1996) combined the pipeline frameworks of Berryman and Oakes to include four aspects of students' experiences in the STEM pipeline: achievement, access, attitude, and activities. In the two decades since, these four aspects of students' experiences have been studied extensively, with clear themes emerging from much of the research regarding how and at what point in a students' K-12 trajectory the experiences may be influenced. Due to the predominant focus of the extant STEM literature on achievement and attitudes, this meta-analysis focuses on these two aspects of the STEM pipeline.

Students' STEM Achievement

Middle grades students' STEM achievement, more than any aspect of STEM pipeline experiences, is a strong predictor of student matriculation through the STEM pipeline to an undergraduate STEM major. The findings of a national survey of U. S. undergraduate students enrolled in introductory courses (n = 6,882) revealed that students who reported higher middle school mathematics grades had 1.5 times higher odds per letter grade to report a STEM career interest at the university level (Dabney et al., 2012). The National Education Longitudinal Study of 1988 (*n* = 4,700) that followed students for twelve years from eighth grade through postsecondary found that students who earned higher scores on eighth-grade mathematics and science assessments were more likely to complete STEM degrees (Maltese & Tai, 2011). For the present study, achievement was defined as any measure of changes in student knowledge of STEM content or processes. In order to retain the greatest number of studies, all types of measures of achievement were included.

Students' STEM-related Attitudes

In addition to studies of students' STEM achievement, students' attitudes towards STEM subjects and careers has also been the subject of a great deal of research. Berryman (1983) and Oakes (1990) were among the first to address the importance of students not only being capable in STEM subjects, but choosing to engage in further study. Most studies of students' STEM attitudes focus on interactions between students' interests and motivation in STEM subjects and the intermediate and/or long-term impacts on students' choice, or intended choice, of choosing STEM careers. Many of the studies are strongly influenced both by social-cognitive theory (Bandura, 2001) and expectancy-value theory (Eccles & Wigfield, 2002), with a focus on the development of students' STEM identities and how they value STEM subject matter and experiences. A longitudinal study tracking middle school students (*n* = 3,359) from eighth grade to age 30 showed the strong connection between students STEM identity in the middle grades and their likelihood to have a career in STEM. Researchers found that students who were interested in a science career in middle school were between 1.9 and 3.4 times more likely to have a science-related career (Dabney et al., 2012). Another study of the science self-perceptions of eighth graders (n = 493) at a diverse urban middle school in California found that students' self-perceptions of their STEM ability was correlated to their perceived value of science and also predicted career interest in science (Aschbacher et al., 2014). For the present study, attitude was broadly defined as any measure of changes in student affect towards STEM subjects or careers; including motivation, efficacy, affinity for STEM, and perceived importance of STEM.

The present study adds to prior meta-analytic work on the effectiveness of STEM interventions, with a specific focus on students in grades five through eight. Though recent meta-analyses of STEM achievement and attitudes (An, 2013; Cheung & Slavin, 2013; Scott et al., 2007) also examined the effectiveness of STEM education across K-12, the present meta-analysis focuses specifically on studies conducted with middle grades students, thus providing a detailed look at the impact of interventions during a critical time in the development of students' identities and achievement in STEM.

The present study addresses the following research questions:

1. What teacher instructional practices are most effective in the development of middle grades students' STEM achievement and attitudes?
2. To what extent do factors such as STEM subject, grade level, school type, SES, and gender significantly moderate the effect of instructional practices on the middle grades students' STEM achievement and attitudes?

Method

Meta-analysis was used in this study due to the need to summarize findings across grade levels, content areas, and outcomes and in order to generate a meaningful comparison of the both the magnitude and direction of the impact of STEM interventions across studies (Lipsey & Wilson, 2000). The critical components of the design were acquisition of studies, establishing criteria for study selection, coding of studies, and the computation of effect size statistics.

Acquisition of Studies

The literature search utilized the Texas A&M University Libraries online search tool to search peer-reviewed journal articles published between 2007 and 2017. The initial list of essential search terms was generated from both the research questions and a broad preliminary literature search, with consultation from a library sciences expert as to which terms were likely to yield the most comprehensive results. The final list of search terms was used in a key word search of three journal databases: ERIC Ebsco, Education Source, and Scopus. In addition, the Tables of Contents of peer-reviewed journals relevant to topics of K-12 education (e.g., *American Education Research Journal*, *Journal of Research on Technology in Education*, *Journal of Research in Mathematics Education*, *Research in Middle Level Education Quarterly*, *Journal of Research in Science Teaching*) were hand-searched. A search of relevant citations in the reference lists of retained articles was also conducted. A complete list of search terms is provided in Appendix A. The Rayyan web-based software for systematic reviews (Ouzzani, Hammady, Fedorowicz, & Elmagarmid, 2016) was used to filter and track search results in study inclusion.

Criteria for Selection of Studies

Each of the studies retained in the initial search were coded by three independent, trained coders using the following criteria: (a) published in the last 10 years (2007-2017), (b) focused on middle-grades students (grades five through eight), (c) included findings relevant to increasing students' STEM achievement or attitude, (d) reported empirical data, and (e) related to the core STEM subject areas of science, technology, engineering, or mathematics. Studies that did not meet initial coding criteria were excluded from the study. Each retained study was coded for general study information, sample characteristics, intervention type, research design, and statistical methods. In order to determine inter-coder reliability, a 10% random sample of articles were re-coded by all coders. Inter-coder reliability across the five articles was 93.1%. The present study adheres to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standards (Moher, Liberati, Tetzlaff, & Altman, 2009). Figure 2.1 provides a PRISMA Flow Diagram illustrating the study inclusion process.

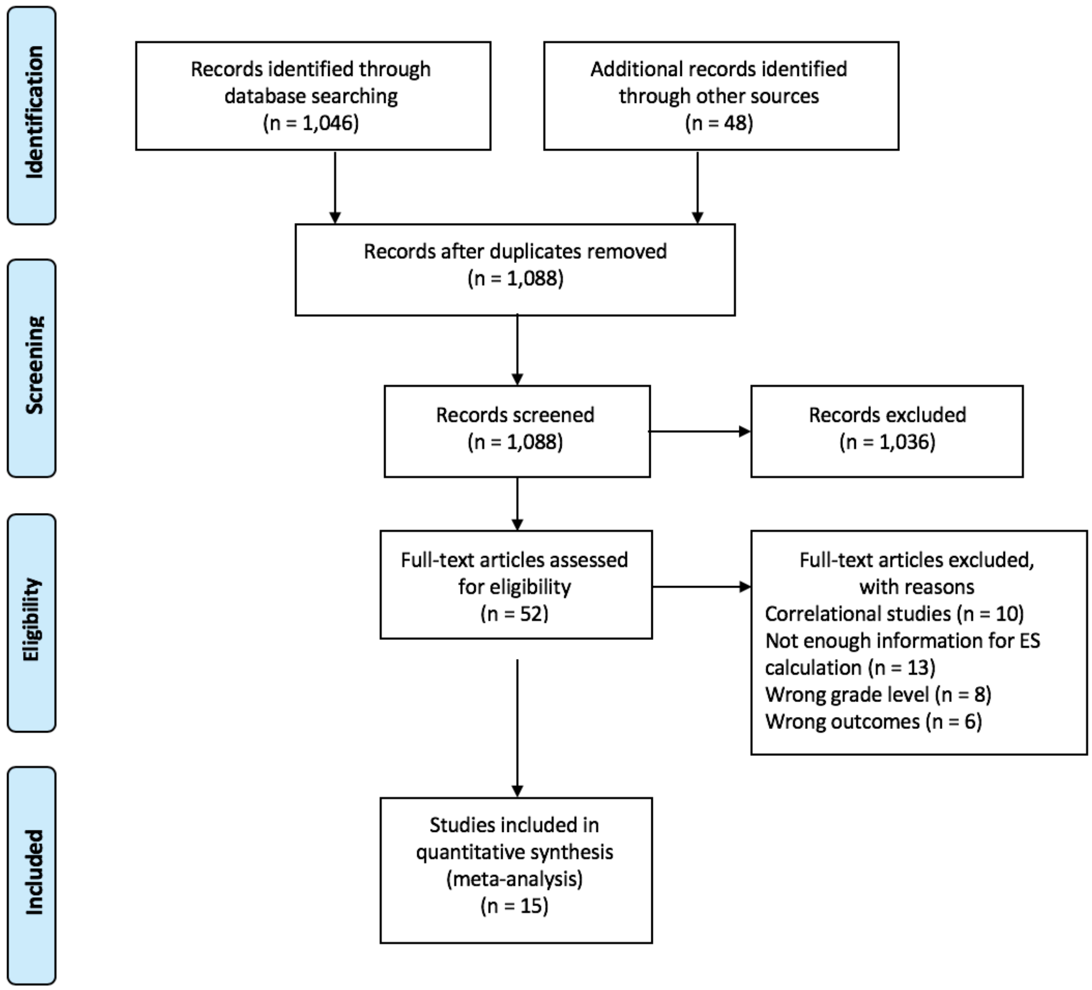


Figure 2.1 *PRISMA Flow Diagram of Study Inclusion.*

Coding of Studies

Articles that were included in the final meta-analytic synthesis (n = 15) included all quantitative studies for which effect size measures could be calculated based on one or more of the following characteristics: (1) an intervention study of contributing factors, (2) a clear control group or pre/post data for one group, and (3a) means and standard deviations for the control and treatment groups or (3b) a calculated effect size with treatment group and control group sample sizes. Correlational studies were not included in the present meta-analysis.

Intercoder objectivity. Each of the studies retained in the initial search was coded by three independent coders using the following criteria: (a) published in the last 10 years (2007-2017), (b) focused on middle-grades students (grades five through eight), (c) included findings relevant to increasing the quality of student STEM experiences, (d) reported empirical data, and (e) related to a core STEM subject area (science, technology, engineering, or mathematics). Studies that did not meet initial coding criteria were excluded from the study. Additionally, each retained study was coded for publication year, intervention characteristics, research design, sample characteristics, and statistical methods. A full code sheet can be found in Appendix B. Additionally, a list of studies included in the final analysis can be found in Appendix C.

Data Analysis

An effect size was calculated for each achievement or attitude measure reported in the included studies. Comprehensive Meta-Analysis Version 3 (Comprehensive Meta-Analysis, 2014) was used in the calculation of effect size measures, determination of publication bias, and the calculation of meta-regression statistics. Hedge's *g* effect sizes were calculated for relevant outcomes from each study, as they provide a less biased measure when comparing studies of differing types and sample sizes (Borenstein, Hedges, Higgins, & Rothstein, 2009). Due to the inclusion of different types of study designs and populations, both fixed and random effects models were used for effect size calculations in order to determine the best estimate of effect size. It was hypothesized that random effects models would be a better estimate of effect size due to the inclusion of studies not equivalent in sample size, population, or method. In addition to the comparison of fixed and random effects models, forest plots were used to examine potential publication bias in the included studies.

Results

Research question one focused on determining the effect of STEM instructional practices on middle grades students' STEM achievement and attitudes. The results of the fixed and random effects models for both outcomes combined and each outcome individually are summarized in Table 2.1. The random-effects model was selected for interpretation for two reasons: (1) the three heterogeneity statistics (Q-statistic, I^2 index, and Tau-squared (τ^2) indicated effect-size heterogeneity, and (2) a random-effects model is more generalizable in this case as it accounts for variability in sample size and study design. The random effects model shows that the STEM interventions in the included studies had a small, positive effect size (Hedge's $g = 0.424$) across all ES ($n = 116$) for both students' STEM achievement and attitude. On average, students involved in STEM interventions performed 0.424 standard deviations higher, on average, than students in the control group in experimental studies or prior to the intervention in pre/post studies. The overall average effect was statistically significant ($p < .001$). The random effects model for the achievement ES only ($n = 58$) shows a slightly higher effect, with a statistically significant moderate effect (Hedge's $g = 0.608$, $p < .001$) for achievement. In contrast, the effect for across all attitude measures was small, but still statistically significant (Hedge's $g = 0.096$, $p = 004$).

Table 2.1

Summary of Fixed- and Random Effects Models

All Effects (Achievement and Attitude)	
Fixed-effects Model	Random-effects Model
$\hat{\theta}$ = .244	$\hat{\theta}$ = .424
CI = [.231, .258]	CI = [.348, .499]
Q(115) = 2203.553, p < .001	Q(115) = 2203.553, p < .001
	I^2 = 94.78%
	$\hat{\tau}^2$ = .111
STEM Achievement	
Fixed-effects Model	Random-effects Model
$\hat{\theta}$ = .276	$\hat{\theta}$ = .608
CI = [.261, .290]	CI = [.509, .706]
Q(57) = 1,972.808, p < .001	Q(57) = 1,972.808, p < .001
	I^2 = 97.11%
	$\hat{\tau}^2$ = .121

Table 2.1 Continued

STEM Attitude	
Fixed-effects Model	Random-effects Model
$\hat{\theta}$ = .052	$\hat{\theta}$ = .096
CI = [.015, .089]	CI = [.030, .163]
Q(57) = 109.399, p < .001	Q(57) = 109.399, p < .001
	I^2 = 47.90%
	$\hat{\tau}^2$ = .020

Figure 2.2 displays the forest plot for all effect sizes in the random effects model.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	Model
Ball et al 1	0.263	1.144	1.310	-1.980	2.506	0.230	0.818	Attitude
Ball et al 10	-0.084	0.938	0.879	-1.922	1.754	-0.089	0.929	Attitude
Ball et al 11	0.017	1.184	1.402	-2.304	2.339	0.015	0.988	Attitude
Ball et al 12	0.051	1.186	1.407	-2.274	2.376	0.043	0.966	Attitude
Ball et al 13	0.106	0.955	0.912	-1.766	1.978	0.111	0.912	Attitude
Ball et al 14	0.369	1.026	1.052	-1.641	2.380	0.360	0.719	Attitude
Ball et al 15	0.038	0.950	0.902	-1.823	1.900	0.040	0.968	Attitude
Ball et al 16	-0.077	0.952	0.906	-1.943	1.789	-0.081	0.936	Attitude
Ball et al 17	-0.074	0.952	0.906	-1.939	1.792	-0.077	0.938	Attitude
Ball et al 18	0.019	0.949	0.901	-1.841	1.879	0.020	0.984	Attitude
Ball et al 19	0.058	1.019	1.039	-1.939	2.056	0.057	0.954	Attitude
Ball et al 2	0.069	1.065	1.135	-2.019	2.156	0.064	0.949	Attitude
Ball et al 20	-0.003	1.017	1.034	-1.996	1.990	-0.003	0.998	Attitude
Ball et al 21	0.054	1.019	1.038	-1.943	2.051	0.053	0.958	Attitude
Ball et al 22	0.026	1.017	1.035	-1.968	2.020	0.026	0.979	Attitude
Ball et al 23	0.026	1.017	1.035	-1.968	2.020	0.026	0.979	Attitude
Ball et al 24	0.017	1.017	1.034	-1.976	2.011	0.017	0.986	Attitude
Ball et al 25	0.051	1.086	1.179	-2.077	2.179	0.047	0.963	Attitude
Ball et al 26	0.164	1.104	1.220	-2.001	2.328	0.148	0.882	Attitude
Ball et al 27	0.164	1.104	1.220	-2.001	2.328	0.148	0.882	Attitude
Ball et al 28	0.028	1.084	1.176	-2.097	2.153	0.025	0.980	Attitude
Ball et al 29	-0.009	1.084	1.174	-2.133	2.115	-0.008	0.994	Attitude
Ball et al 3	0.035	1.061	1.126	-2.044	2.115	0.033	0.973	Attitude
Ball et al 30	0.033	1.084	1.176	-2.092	2.159	0.031	0.975	Attitude
Ball et al 4	0.027	1.060	1.124	-2.051	2.106	0.026	0.979	Attitude
Ball et al 5	0.032	1.061	1.125	-2.047	2.111	0.030	0.976	Attitude
Ball et al 6	-0.002	1.059	1.122	-2.079	2.074	-0.002	0.998	Attitude
Ball et al 7	0.235	0.990	0.979	-1.705	2.174	0.237	0.813	Attitude
Ball et al 8	0.204	0.975	0.951	-1.707	2.116	0.210	0.834	Attitude
Ball et al 9	0.070	0.936	0.875	-1.764	1.903	0.075	0.940	Attitude
Brown et al	0.581	0.105	0.011	0.375	0.787	5.527	0.000	Attitude
Chen & Howard 1	0.000	0.084	0.007	-0.165	0.165	0.000	1.000	Attitude
Chen & Howard 10	0.193	0.269	0.072	-0.334	0.720	0.718	0.473	Achievement
Chen & Howard 11	1.250	0.169	0.028	0.920	1.581	7.419	0.000	Achievement
Chen & Howard 12	0.033	0.290	0.084	-0.535	0.602	0.115	0.908	Achievement
Chen & Howard 13	1.405	0.329	0.108	0.761	2.049	4.273	0.000	Achievement
Chen & Howard 14	0.255	0.275	0.075	-0.283	0.793	0.929	0.353	Achievement
Chen & Howard 2	0.205	0.085	0.007	0.039	0.371	2.418	0.016	Attitude
Chen & Howard 3	-0.044	0.084	0.007	-0.209	0.120	-0.530	0.596	Attitude
Chen & Howard 4	0.048	0.084	0.007	-0.117	0.212	0.567	0.571	Attitude
Chen & Howard 5	0.022	0.084	0.007	-0.143	0.188	0.262	0.793	Attitude
Chen & Howard 6	-0.105	0.085	0.007	-0.272	0.062	-1.234	0.217	Attitude
Chen & Howard 7	0.026	0.085	0.007	-0.139	0.192	0.311	0.756	Attitude
Chen & Howard 8	0.003	0.310	0.096	-0.606	0.611	0.008	0.993	Achievement
Chen & Howard 9	0.939	0.403	0.162	0.150	1.728	2.332	0.020	Achievement
Hayden et al 1	-0.629	0.310	0.096	-1.237	-0.020	-2.025	0.043	Achievement
Hayden et al 2	0.414	0.293	0.086	-0.160	0.989	1.414	0.157	Achievement
Hayden et al 3	0.525	0.215	0.046	0.104	0.946	2.442	0.015	Achievement
Hayden et al 4	0.402	0.287	0.082	-0.160	0.964	1.401	0.161	Achievement
Hayden et al 5	0.641	0.318	0.101	0.018	1.265	2.015	0.044	Achievement
Hseih et al 1	0.012	0.065	0.004	-0.116	0.140	0.185	0.853	Attitude
Hseih et al 2	-0.246	0.064	0.004	-0.372	-0.121	-3.844	0.000	Attitude
Hseih et al 3	-0.164	0.064	0.004	-0.291	-0.038	-2.551	0.011	Attitude
Hseih et al 4	0.164	0.064	0.004	0.038	0.291	2.547	0.011	Attitude
Hseih et al 5	1.348	0.069	0.005	1.212	1.484	19.446	0.000	Achievement

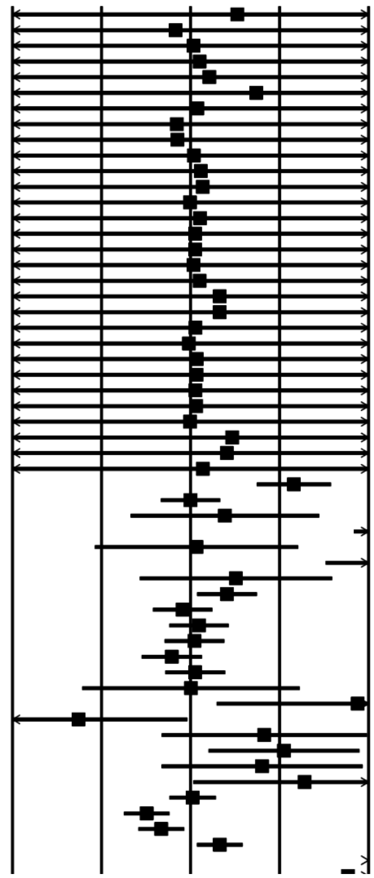
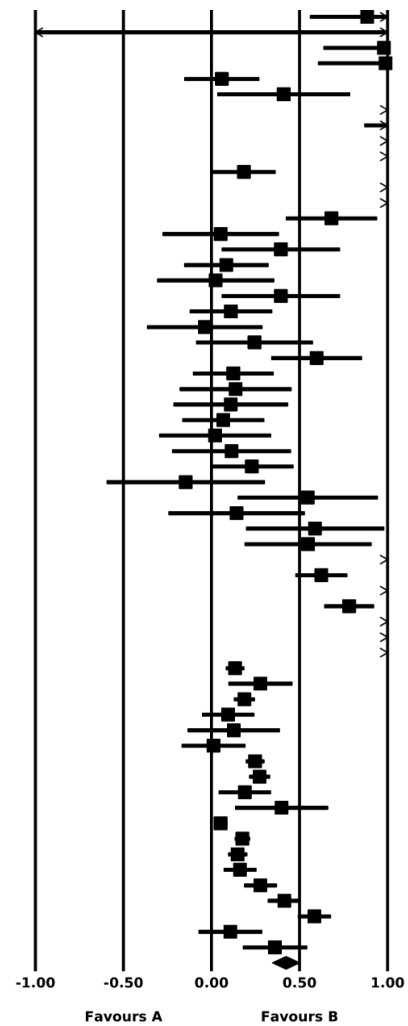


Figure 2.2. *Forest Plot for Random Effects Model.*

Figure 2.2 Continued

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	
Jackson et al 1	0.885	0.165	0.027	0.562	1.208	5.372	0.000	Attitude
Jackson et al 2	2.476	1.873	3.509	-1.195	6.148	1.322	0.186	Achievement
Jackson et al 3	0.980	0.174	0.030	0.638	1.321	5.626	0.000	Achievement
Jackson et al 4	0.989	0.194	0.038	0.608	1.369	5.094	0.000	Achievement
Johnson, Kahle, & Fargo 1	0.059	0.107	0.011	-0.151	0.269	0.550	0.582	Achievement
Johnson, Kahle, & Fargo 2	0.411	0.191	0.036	0.037	0.784	2.153	0.031	Achievement
Johnson, Kahle, & Fargo 3	1.251	0.118	0.014	1.021	1.482	10.640	0.000	Achievement
Johnson, Kahle, & Fargo 4	1.246	0.192	0.037	0.871	1.622	6.503	0.000	Achievement
Johnson, Kahle, & Fargo 5	1.502	0.122	0.015	1.263	1.740	12.347	0.000	Achievement
Johnson, Kahle, & Fargo 6	1.830	0.223	0.050	1.393	2.267	8.204	0.000	Achievement
Johnson, Kahle, & Fargo 7	0.185	0.090	0.008	0.008	0.362	2.044	0.041	Achievement
Johnson, Kahle, & Fargo 8	1.327	0.107	0.012	1.117	1.538	12.359	0.000	Achievement
Johnson, Kahle, & Fargo 9	1.537	0.110	0.012	1.321	1.753	13.940	0.000	Achievement
Knezek et al 1	0.681	0.131	0.017	0.425	0.938	5.211	0.000	Achievement
Knezek et al 10	0.052	0.167	0.028	-0.275	0.380	0.314	0.753	Attitude
Knezek et al 11	0.394	0.170	0.029	0.061	0.727	2.322	0.020	Attitude
Knezek et al 12	0.085	0.121	0.015	-0.152	0.321	0.702	0.483	Attitude
Knezek et al 13	0.024	0.168	0.028	-0.306	0.353	0.141	0.888	Attitude
Knezek et al 14	0.394	0.170	0.029	0.061	0.727	2.322	0.020	Attitude
Knezek et al 15	0.111	0.118	0.014	-0.121	0.342	0.938	0.348	Attitude
Knezek et al 16	-0.038	0.166	0.027	-0.363	0.287	-0.229	0.819	Attitude
Knezek et al 17	0.244	0.168	0.028	-0.084	0.573	1.459	0.145	Attitude
Knezek et al 2	0.597	0.130	0.017	0.343	0.852	4.597	0.000	Achievement
Knezek et al 3	0.124	0.115	0.013	-0.102	0.350	1.074	0.283	Attitude
Knezek et al 4	0.137	0.161	0.026	-0.178	0.452	0.853	0.393	Attitude
Knezek et al 5	0.110	0.165	0.027	-0.213	0.433	0.665	0.506	Attitude
Knezek et al 6	0.067	0.117	0.014	-0.162	0.297	0.574	0.566	Attitude
Knezek et al 7	0.021	0.160	0.026	-0.293	0.335	0.131	0.896	Attitude
Knezek et al 8	0.114	0.171	0.029	-0.221	0.449	0.668	0.504	Attitude
Knezek et al 9	0.229	0.119	0.014	-0.004	0.463	1.925	0.054	Attitude
Lara-Alecio et al 1	-0.146	0.228	0.052	-0.592	0.300	-0.643	0.520	Achievement
Lara-Alecio et al 2	0.547	0.201	0.041	0.152	0.942	2.715	0.007	Achievement
Lara-Alecio et al 3	0.143	0.196	0.038	-0.241	0.527	0.729	0.466	Achievement
Lara-Alecio et al 4	0.589	0.199	0.040	0.199	0.978	2.961	0.003	Achievement
Lara-Alecio et al 5	0.548	0.183	0.033	0.191	0.906	3.004	0.003	Achievement
Lee et al 1	1.537	0.082	0.007	1.377	1.698	18.756	0.000	Achievement
Lee et al 2	0.624	0.074	0.005	0.480	0.769	8.465	0.000	Achievement
Lee et al 3	1.599	0.078	0.006	1.446	1.752	20.448	0.000	Achievement
Lee et al 4	0.782	0.071	0.005	0.644	0.920	11.073	0.000	Achievement
Liu et al 1	1.477	0.208	0.043	1.069	1.886	7.089	0.000	Achievement
Liu et al 2	1.783	0.180	0.032	1.430	2.135	9.920	0.000	Achievement
Liu et al 3	1.641	0.136	0.018	1.374	1.907	12.065	0.000	Achievement
Llosa et al 1	0.134	0.025	0.001	0.084	0.184	5.285	0.000	Achievement
Llosa et al 10	0.278	0.091	0.008	0.099	0.457	3.047	0.002	Achievement
Llosa et al 2	0.187	0.029	0.001	0.130	0.244	6.441	0.000	Achievement
Llosa et al 3	0.095	0.074	0.006	-0.050	0.240	1.281	0.200	Achievement
Llosa et al 4	0.126	0.132	0.017	-0.133	0.385	0.957	0.339	Achievement
Llosa et al 5	0.012	0.091	0.008	-0.167	0.190	0.127	0.899	Achievement
Llosa et al 6	0.248	0.025	0.001	0.198	0.297	9.714	0.000	Achievement
Llosa et al 7	0.273	0.029	0.001	0.216	0.330	9.379	0.000	Achievement
Llosa et al 8	0.190	0.074	0.006	0.044	0.335	2.551	0.011	Achievement
Llosa et al 9	0.399	0.133	0.018	0.137	0.660	2.990	0.003	Achievement
Marshall et al 1	0.052	0.017	0.000	0.018	0.086	3.015	0.003	Achievement
Marshall et al 2	0.176	0.021	0.000	0.135	0.217	8.387	0.000	Achievement
Marshall et al 3	0.149	0.026	0.001	0.097	0.200	5.691	0.000	Achievement
Scott et al 1	0.163	0.046	0.002	0.072	0.253	3.533	0.000	Achievement
Scott et al 2	0.278	0.046	0.002	0.188	0.368	6.027	0.000	Achievement
Scott et al 3	0.414	0.046	0.002	0.323	0.504	8.969	0.000	Achievement
Scott et al 4	0.585	0.046	0.002	0.494	0.675	12.663	0.000	Achievement
Zweip & Straits 1	0.107	0.091	0.008	-0.071	0.285	1.177	0.239	Achievement
Zweip & Straits 2	0.361	0.092	0.008	0.182	0.540	3.943	0.000	Achievement
Random	0.424	0.039	0.001	0.348	0.499	10.979	0.000	

Model; Statistics for each study



In order to determine if selection bias was present in the included studies, funnel plots were created and analyzed for the presence of outliers. Asymmetrically-shaped funnels and effect sizes falling outside of the funnel indicate that selection bias may be present in the study (Anzures-Cabrera & Higgins, 2010). Figure 2.3 displays the funnel plot for the random effects model. The presence of quite a few outliers outside of the funnel indicated that publication bias may be present in the study.

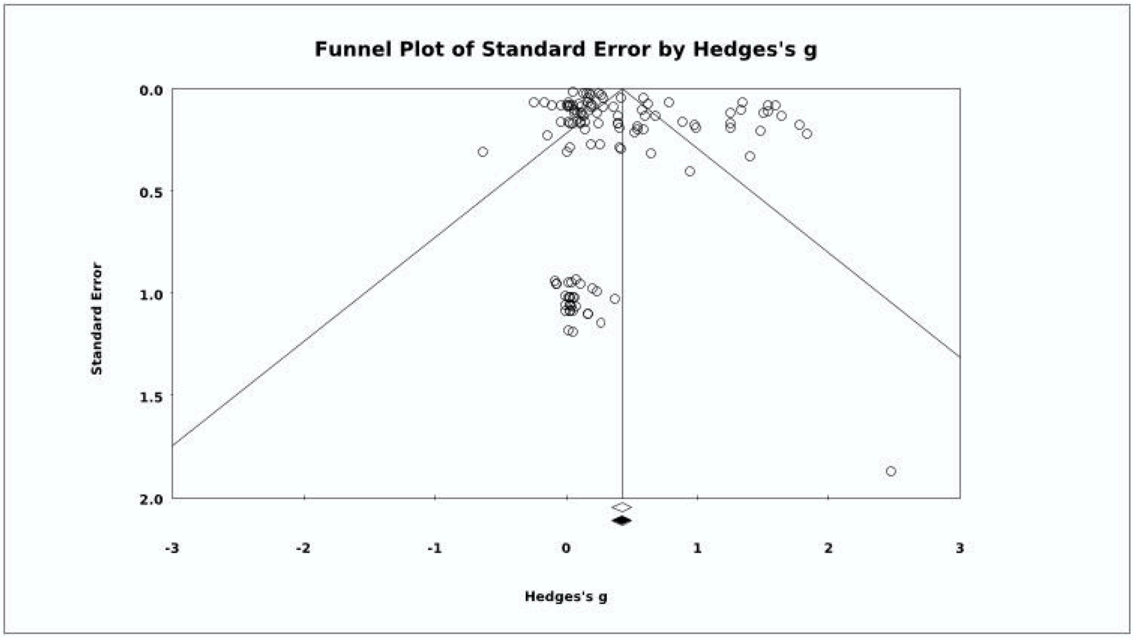


Figure 2.3. *Funnel Plot for Random Effects Model.*

In addition to the funnel plot, Duval and Tweedie's (2000) trim and fill approach was used to determine studies to the left and right of the mean that were potentially influencing effect size estimates. The trim and fill approach indicated zero missing studies to the left of the mean and 26 missing studies to the right of the mean. Using the one study removed method that estimates the average effect size if effect sizes are removed one at a time, effect size estimates ranged from .408 to .433. Based on these findings, it appears that there is some selection bias present in the study.

Research question two focused on focused on the extent to which factors including publication year, intervention type, grade level, duration of intervention, student status as an English language learner (ELL), student ethnicity, test type, student gender, and content area significantly moderated the effect of instructional practices on the middle grades students' STEM achievement and attitudes. A meta-regression using restricted-maximum likelihood estimation (REML) was conducted on the combined outcomes, as well as on each outcome individually, to determine if the set of moderators explained a significant amount of variance in the study impacts.

STEM Attitude and Achievement

The R^2 of the full model including both achievement and attitude was 0.71, indicating that the moderators explained 71% of the variability in the effect sizes across studies. Four variables were determined to be significant moderators of overall effect size for both outcomes, including grade level, duration of the intervention, test type, and student gender. Compared to the reference group (grade 5), grade level has a positive effect for grades six, seven, and eight. With regard to duration of intervention, studies with a duration of greater than one year had a statistically significant positive effect compared to studies with a duration of less than one year. The use of previously validated tests to measure outcomes had a statistically significant negative effect on attitude and achievement compared to studies utilizing standardized measures. Finally, being female was statistically positively associated with STEM achievement and attitude outcomes. Results of the moderator analyses for the full model are displayed in Table 2.2.

Table 2.2

Moderator Analysis for Students' STEM Achievement and Attitudes

Moderator	Estimate	SE	z-value	p-value
Publication year	-0.054	0.031	-1.74	0.082
Intervention type				
Inquiry and technology	-0.179	0.215	-0.83	0.405
Inquiry and English language/vocabulary development	-0.386	0.458	-0.84	0.399
Technology only	0.232	0.332	0.7	0.485
Test type				
Research-created	0.227	0.150	1.51	0.131
Teacher/district-created	0.245	0.356	0.69	0.491
Previously-validated instrument	-0.565	0.211	-2.68	0.007**
Grade level				
Sixth	1.259	0.335	3.75	0.000***
Seventh	1.559	0.360	4.33	0.000***
Eighth	1.700	0.359	4.73	0.000***
Multiple grade levels	0.545	0.282	1.94	0.053
Duration of intervention				
One year	0.159	0.153	1.04	0.300

Table 2.2 Continued

Moderator	Estimate	SE	z-value	p-value
Two years	0.625	0.268	2.33	0.012*
Three years	0.833	0.243	3.420	0.001**
Four years	0.913	0.315	2.900	0.004**
Five years	1.055	0.392	2.690	0.007**
Student ELL status				
ELL	-0.185	0.170	-1.090	0.275
Not reported	-0.284	0.186	-1.530	0.127
Student ethnicity				
African American	0.080	0.196	0.410	0.683
Hispanic	0.625	0.266	2.350	0.019
Student gender				
Female only	0.336	0.124	2.710	0.007**
Female majority sample	0.294	0.108	2.710	0.007**
Content area				
Science	0.076	0.199	0.380	0.703
Technology	0.082	0.215	0.380	0.703
Multiple STEM subjects	-0.004	0.228	-0.020	0.985
Engineering	0.050	0.237	0.210	0.833
Intercept	108.535	62.902	1.730	0.084

Note. $*p < .05$, $**p < .01$, $***p < .001$.

STEM Achievement

The R^2 of the model including achievement effects was 0.78, indicating that the moderators explained 78% of the variability in the achievement effect sizes across studies. Five variables were determined to be significant moderators of overall effect size for STEM achievement, including intervention type, grade level, duration of the intervention, test type, and student gender. Compared to inquiry-only interventions, interventions including inquiry along with a technology component or an English-language development component had statistically significant positive effects on students' STEM achievement. Interventions with a technology-only component had a statistically significant negative effect on students' STEM achievement compared to inquiry-only interventions. Grade level had a significant positive effect for sixth-grade students' STEM achievement, as well as for studies of mixed grade level groups. With regard to duration of intervention, studies lasting longer than one year, with the exception of studies with a duration of greater than four years had a statistically significant negative effect on achievement compared to studies with a duration of less than one year. Compared to standardized tests, the use of previously validated and researcher-created tests to measure students' achievement outcomes had a statistically significant positive effect on achievement outcomes. Finally, effects including both males and females were statistically and positively associated with STEM achievement compared to studies focused on males only. Results of the moderator analyses for achievement are displayed in Table 2.3.

Table 2.3

Moderator Analysis for Students' STEM Achievement

Moderator	Estimate	SE	z-value	p-value
Publication year	0.081	0.061	1.310	0.189
Intervention type				
Inquiry and technology	0.739	0.352	2.100	0.036*
Inquiry and English language/vocabulary development	1.597	0.676	2.360	0.018*
Technology only	-3.838	1.037	-3.700	0.000***
Test type				
Research-created	0.070	0.157	0.450	0.655
Teacher/district-created	0.866	0.440	1.970	0.049*
Previously-validated instrument	1.744	0.622	2.800	0.005**
Grade level				
Sixth	1.026	0.395	2.590	0.010**
Seventh	0.142	0.591	0.240	0.810
Eighth	0.895	0.527	1.700	0.090
Multiple grade levels	1.075	0.530	2.030	0.043*
Duration of intervention				
One year	-2.041	0.642	-3.180	0.002**
Two years	-1.726	0.707	-2.440	0.015*

Table 2.3 Continued

Moderator	Estimate	SE	z-value	p-value
Three years	-1.571	0.705	-2.230	0.026*
Four years	-1.458	0.730	-2.000	0.046*
Five years	-1.285	0.763	-1.680	0.092
Student ELL status				
ELL	-0.016	0.187	-0.090	0.931
Not reported	0.049	0.211	0.230	0.817
Student ethnicity				
African American	0.156	0.219	0.710	0.475
Hispanic	-1.048	0.515	-2.030	0.042*
Student gender				
Females only	0.308	0.330	0.930	0.350
Female majority sample	0.834	0.230	3.630	0.000***
Intercept	-161.247	123.051	-1.310	0.190

Note. $*p < .05$, $**p < .01$, $***p < .001$.

STEM Attitude

The R^2 of the model including attitude effects only was 0.96, indicating that the moderators explained 96% of the variability in the attitude effect sizes across studies. One variable, student gender, was determined to be a significant moderator of overall effect size for students' STEM attitudes. Compared to males, the included studies had a positive and significant impact on females. Due to collinearity of moderators due to a smaller number of studies with attitude as a dependent measure, the moderator analysis for attitude contained fewer moderators overall. Results of the moderator analyses for attitude are displayed in Table 2.4.

Table 2.4

Moderator Analysis for Students' STEM Attitudes

Moderator	Estimate	SE	z-value	p-value
Publication year	-0.009	0.034	-0.270	0.791
Test type				
Research-created	-0.831	0.515	-1.620	0.106
Previously-validated instrument	-1.343	0.758	-1.770	0.076
Grade level				
Sixth	0.453	0.536	0.850	0.397
Multiple grade levels	0.106	0.503	0.210	0.833
Student ethnicity				
African American	0.051	0.569	0.090	0.928
Hispanic	0.473	0.244	1.940	0.052
Student gender				
Female	0.248	0.056	4.460	0.000***
Not reported	0.101	0.082	1.240	0.215

Table 2.4 Continued

Moderator	Estimate	SE	z-value	p-value
Content area				
Science	-0.038	0.118	-0.320	0.747
Technology	0.108	0.118	0.920	0.360
Multiple STEM subjects	0.014	0.119	0.120	0.904
Engineering	0.037	0.122	0.310	0.758
Intercept	18.646	67.436	0.280	0.782

Note. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Discussion

The present meta-analysis examined the impact of STEM interventions on middle-grade students' achievement and attitudes in STEM subjects, as well as what factors moderate the impact of STEM interventions. On average, students involved in STEM interventions performed 0.424 standard deviations higher than students in the control group in experimental studies or prior to the intervention in pre/post studies. The overall average effect was statistically significant ($p < .001$), with a slightly higher effect for achievement (Hedge's $g = 0.608$, $p < .001$). In contrast, the average across all attitude measures was small, but still statistically significant (Hedge's $g = 0.096$, $p =$ 004), with students' STEM attitudes 0.096 standard deviations higher than non-intervention students or prior to an intervention.

The overall impact of STEM interventions aligns with another recent review of the impact of STEM programs on similar constructs. An (2013) found small positive impacts for students' engagement, or attitude (0.346), and capability, or achievement (0.454). The impact of interventions on middle grades students' STEM achievement mirror the results of Schroeder et al's (2007) meta-analysis of the impact of teaching strategies on K-12 students' science achievement. Schroeder et al found an overall moderate and significant effect for teaching strategies of 0.66 across all studies in grades K-8.

Research question two examined the degree to which factors such as grade level and duration of intervention moderated the impact of study results. Grade level had a positive and significant impact on both attitude and achievement for grades six through eight when achievement and attitude effects were combined, as well as a positive and significant impact on grade eight in achievement. There were no significant grade level impacts on attitude. The impact of grade level on middle-grades students' STEM outcomes in previous reviews have somewhat conflicting findings. Similar to the present study, Schroeder et al (2007) found that the impact of teaching strategies on students' science achievement increased as students entered higher grades. However, in contrast to the present study's findings and to Schroeder et al., other meta-analyses of both STEM achievement and attitude found that as students grade level increases, the impact of interventions decreases (An, 2013; Cheung & Slavin, 2013).

Study duration appeared to have differential impacts on study effects, with an overall significant and positive impact for duration of greater than one year. However, when viewed separately, study duration of longer than one year appears to have a negative impact on achievement. This could be explained by students' achievement measures regressing to the mean with repeated testing, as well as a testing validity threat that is introduced when measures are repeated over the course of multiple years in a study. Schroeder et al (2007) also found a negative, though non-significant, impact of longer study duration on students' science achievement. There were no significant study duration impacts on attitude.

Similar to study duration, test type (researcher-created, teacher/district-created, standardized, or previously validated for use in another study) had differential impacts on study effects, with an overall significant and negative impact for the use of instruments previously used in another study compared to the use of standardized instruments. However, the use of previously validated or researcher-created instruments had a positive impact on achievement. These findings stand in contrast to those of Schroeder et al (2007) who found significant negative impacts for non-standardized test types. There were no significant test type impacts on attitude.

Two remaining moderators, student gender and intervention type, had significant impacts on study effects. Across all studies, being female had significant and positive impacts on study outcomes compared to the male reference group. This finding is encouraging, given the fact that previous studies of females' achievement and attitudes in STEM have shown that female attitudes tend to be less positive, decline more steadily with age than males, and females have lower perceptions of competence even if they enjoy science (Brotman & Moore 2008; Catsambis 1995, Riegle-Crumb, King, Grodsky, & Muller, 2012).

Finally, intervention types that included combinations of interventions compared to inquiry-only interventions had significant positive impacts on study outcomes, while technology-only interventions had significant negative impacts on achievement. Compared to inquiry-only interventions, interventions including inquiry along with a technology component or an English-language development component had statistically significant positive effects on students' STEM achievement. These findings conflict with previous reviews of students' achievement and attitude in STEM (An, 2013; Cheung & Slavin, 2013) that found mixed, or comprehensive, interventions do not have more positive impacts than either inquiry or technology alone. This conflicting finding may be due to the fact that several of the studies included in the present meta-analysis focused specifically on English-language learners and the addition of an English-language development component may have played a strong role in the positive impacts of comprehensive interventions for ELLs. There were no significant intervention type impacts on attitude.

Study Implications

The present study's findings add to the body of literature on the effectiveness of reform-based STEM instructional interventions in the middle grades. It appears that reform-based STEM instruction involving inquiry, technology, and vocabulary or language development yields a small to moderate effect size, especially for girls and upper-grades students. However, it appears to be more difficult to move the needle on attitude than achievement, as only small effect sizes were found for attitude compared to moderate effect sizes for achievement. In addition, comprehensive interventions seem to work better than isolated interventions in the middle grades.

Regarding duration and measurement, it appears that there is a point of diminishing returns in the measurement of achievement, pointing to potential validity threats with repeated measurement and/or multi-year studies. It is important that researchers and educators carefully consider how, and how often, to best measure outcomes, particularly for achievement. It may be that more frequent formative assessments are more effective in measuring changes in student achievement. In addition, frequent formative assessment would provide education researchers and educators with information as to what adjustments to interventions might result in greater impacts for students. Another assessment consideration is the present study's finding of positive and significant impacts for researcher-created and previously validated instruments to measure achievement in the middle grades over standardized testing. This has important implications for state testing and school accountability measures that solely use standardized testing as a measure of student achievement in STEM.

Limitations of Study

This study, though important in its specific focus on middle grades STEM, has several limitations. Limiting studies to only those addressing middle-grade students resulted in a small total number of studies (*n* = 15). A total of 10 correlational studies that were initially retained were not included in the final meta-analysis due to limitations of the research designs and correlational natures of the data. In addition to a small total number of studies, outliers in the funnel plots, combined with the fact that the present study only used peer-reviewed articles, likely resulted in publication bias in the included studies.

Conclusion

The present meta-analysis examined the impact of STEM interventions on middle grades students' achievement and attitudes in STEM subjects, as well as what factors moderate the impact of STEM interventions. Overall, STEM interventions in the included studies have a positive, moderate impact (Cohen's d = 0.424) on both achievement and attitude of middle grades students. These findings are in line with a synthesis of meta-analytic findings of STEM interventions (Hattie, 2009), that found Cohen's d effect sizes ranging from 0.23 for interventions focused on technology in science to 0.59 for mathematics programs. The present study's findings also provide evidence that STEM interventions may have a greater impact on middle grades students, and female middle grades students in particular, than educational interventions as a whole. Hattie's (2009) meta-synthesis found a small average effect size of 0.08 for both middle school interventions overall and for the impact of gender on achievement, while the present study found larger overall effects for middle grades students (0.434) and for female students (0.308).

The present study's findings and those of other recent reviews of STEM achievement and attitude (An, 2013; Cheung & Slavin, 2013; Schroeder et al., 2007) point to a need for educators and policy makers to carefully evaluate the impact of different types of STEM interventions in their states, districts, and campuses; as well as ways in which students’ STEM outcomes are measured.

CHAPTER III

A MULTILEVEL ANALYSIS OF SCHOOL AND TEACHER FACTORS CONTRIBUTING TO EFFECTIVE STEM PRACTICES IN THE MIDDLE GRADES

As the U. S. has shifted its focus from a skills-based economy to a largely knowledge-based economy, student success in the fields of science, technology, engineering, and mathematics (STEM), as well as in the overlapping areas between these fields, continues to be a topic of great national interest. STEM skills, such as process-based thinking skills including problem ideation and problem solving, persistence, and creativity (Bailey, Kaufman, & Subotic, 2015) are required for over 80% of the 30 fastest growing occupations in the United States (U. S. Bureau of Labor Statistics, Employment Projections Program, 2017). In addition, the analytical skills related to STEM disciplines are in demand across a variety of jobs, including construction, manufacturing, public administration, and management (Rothwell, 2013).

The continued growth of a STEM-focused economy has resulted in an intensified focus on the development of K-12 students' STEM competencies and therefore, on the STEM-related knowledge and practices of K-12 STEM teachers. In fact, teacher quality has been identified as the single most impactful factor in raising student STEM achievement, specifically in mathematics, (Hattie 2009; Rivkin, Hanushek, & Kain 2005) and is increasingly viewed as paramount to increasing numbers of STEM-skilled students matriculating through the STEM pipeline into higher education or directly into the STEM workforce (Carnegie Commission for Mathematics and Science Teaching, 2009; Coble, 2012; National Academy of Sciences, National Academy of Engineering & Institute of Medicine, 2007; NGSS Lead States, 2013; National Research Council, 2012; Wilson, 2013).

The result of increased scrutiny has led to a great deal of examination and standard-setting focused around STEM teachers' instructional practices and STEM student achievement in recent years. However, though national STEM advocacy groups, such as the American Statistical Association, the National Academy of Sciences, National Academy of Engineering, and Institute of Medicine, advocate frameworks focused on STEM improvement, few studies have provided actionable information for campus and district-level administrators on what alterable school- or teacher-level factors impact the degree to which STEM teachers utilize reform-based instruction.

The present study examines the ways in which teaching practices related to the four aspects of students' STEM pipeline experience (achievement, attitudes, access, and activities) are impacted by malleable school and classroom factors. The study focuses on middle grades (grades 5 – 8) students and teachers due to overwhelming evidence in the extant literature that the middle grades is a time when students' attitudes regarding STEM fields and careers are most subject to change (Aschbacher, Ing, & Tsai, 2014; Catsambis, 1995; Christensen & Knezek, 2015; Hinojosa, Rapaport, Jaciw, LiCalsi, & Zacamy, 2016; Ing & Nylund-Gibson, 2017; Nugent, Barker, Welch, Grandgenett, Wu, & Nelson, 2015; Oakes, 1990). Though there is some inconsistency in how middle grades is defined, the present study utilized the U. S. Department of Education's Institute for Education Sciences operationalization of middle grades as grade levels five through eight (U. S. Department of Education, Institute for Education Sciences, National Center for Education Statistics, 2015). The study focuses on determining: (a) the impact of school factors on middle grades STEM teachers' use of effective STEM practices and (b) the impact of teachers' perceptions of their professional learning, feelings of preparedness, and availability of resources on their use of effective STEM practices. Secondary data analysis is utilized to explore teacher and school factors that explain variation in teacher self-reports of effective STEM practices in a nationally representative sample of middle grades STEM teachers surveyed in the National Survey of Science and Mathematics Educators (NSSME, Weis & Banilower, 2014).

A great deal of research has been conducted on one or more aspects of students' middle grades STEM pipeline experiences. However, the majority of studies rely on student-reported data. Few studies provide an account of the degree to which teachers perceive that their schools and classrooms are providing students with the types of experiences found in research to be influential in their matriculation to STEM undergraduate majors and careers, as well as how teacher perceptions of their practices contribute to student experiences and outcomes in STEM.

In addition, much of the current research focuses on factors fixed factors that cannot be manipulated at the school level to directly impact students. For example, Hanson's (1996) consideration of external factors acting on the STEM pipeline was limited mostly to fixed ideas such as family structure, school characteristics, and courses taken. Though Hansen's framework did include an examination of a few malleable factors, such as teacher and student attitudes, missing from the framework were mechanisms that research has more recently identified as impactful in general school improvement efforts (Allensworth, Ponisciak, & Mazzeo, 2009; Brophy, Klein, Portsmore, & Rogers, 2008; Bryk & Schneider, 2002; DeAngelis & Presley, 2011; Maltese & Tai, 2011), such as school culture (i.e., level of collaboration, classroom autonomy), professional learning (i.e., amount and types of professional development in STEM), and specific instructional practices (i.e., classroom discourse, hands-on learning experiences, real world connections).

In 2013, the Institute of Education Sciences, the U. S. Department of Education, and the National Science Foundation emphasized the importance of both alterable and ground-level factors with the establishment of a focus on malleable school and classroom factors as a requirement of all early stage or exploratory research programs seeking federal funding from either organization. The institutions defined malleable factors as, "…factors that are alterable, such as children's behaviors; technologies; education programs; policies; and practices," (2013, p. 12). A recent review of factors impacting Hispanic student success in STEM (Hinojosa et al., 2016) also emphasized the importance of focusing on indicators that are: (a) predictive of student success in STEM, (b) malleable, and (c) actionable at the school or district, rather than state or federal, levels. However, less than a quarter of the studies reviewed focused on school or teacher characteristics and only one study focused on STEM pedagogy. None of the studies reviewed focused on differences in school- or classroom-level predictors between Hispanic and non-Hispanic students. In order for schools and classroom teachers to maximize the effectiveness of ground-level efforts to broaden the STEM pipeline, it is critical to investigate the impact of school culture and teacher professional learning on teachers' use of effective STEM practices.

The purpose of the present study is to explore teacher and school factors that explain variation in teacher self-reports of use of effective STEM practices in a nationally representative sample of middle grades STEM teachers. A secondary purpose is to explore whether there are differences in the use of effective STEM practices by school. The NSSME was chosen over more commonly used nationally-representative surveys (e.g., Schools and Staffing Survey, National Teacher and Principal Survey, Trends in International Mathematics and Science Survey, OECD Teaching and Learning International Survey) due to its focus on U. S. – based K-12 STEM teachers. The analysis centers around teachers' perceptions of reform-based instructional objectives, instructional practices, and use of instructional technology. The following research questions guide the study:

1. What proportion of the variance in teachers' use of effective STEM practices is attributable to school differences? Is there a significant variation among schools in the use of effective STEM practices?
2. What is the effect of malleable school- and teacher-level factors on teachers' use of effective STEM instructional objectives?
3. What is the effect of malleable school- and teacher-level factors on teachers' use of effective STEM instructional practices?
4. What is the effect of malleable school- and teacher-level factors on teachers' use of instructional technology?

Conceptual Framework

The conceptual framework for the present study is adapted from Hansen's (1996) STEM pipeline factors, including achievement, attitude, access, and activities, and also takes into account more recent work on impactful mechanisms for teacher and school change (Allensworth et al., 2009; Brophy et al., 2008; Bryk & Schneider, 2002; DeAngelis & Presley, 2011; Maltese & Tai, 2011), including school culture (i.e., level of collaboration, classroom autonomy), professional learning (i.e., amount and types of professional development in STEM), and specific instructional practices (i.e., classroom discourse, hands-on learning experiences, real world connections). The school- and teacher level predictors are centered around the *Contexts for Teachers' Learning* framework (National Academies of Sciences, Engineering, and Medicine's Committee on Strengthening Science Education, 2015), which is based on a synthesis of research in school- and teacher-level factors that create supportive contexts for STEM teaching and learning. The framework is based on multiple multiple studies of comprehensive school reform efforts (Cohen & Hill, 2000, 2001; Cohen, Raudenbush, & Ball, 2003; Le et al, 2006; Rowan, Corenti, Miller, & Camburn, 2009), including the work of the Chicago Consortium for School Reform (Bryk, Sebring, Allensworth, Suppescu, & Easton, 2010).

The four main aspects of the framework are: (1) professional capacity, (2) coherent instructional guidance, (3) leadership, and (4) time and funding. The overarching purpose of the conceptual framework is to provide a lens through which

K-12 school stakeholders might examine the impact of school- and teacher-level factors on students' STEM experiences in order to increase positive STEM outcomes for students. Figure 3.1 provides an overview of the conceptual framework of the study.

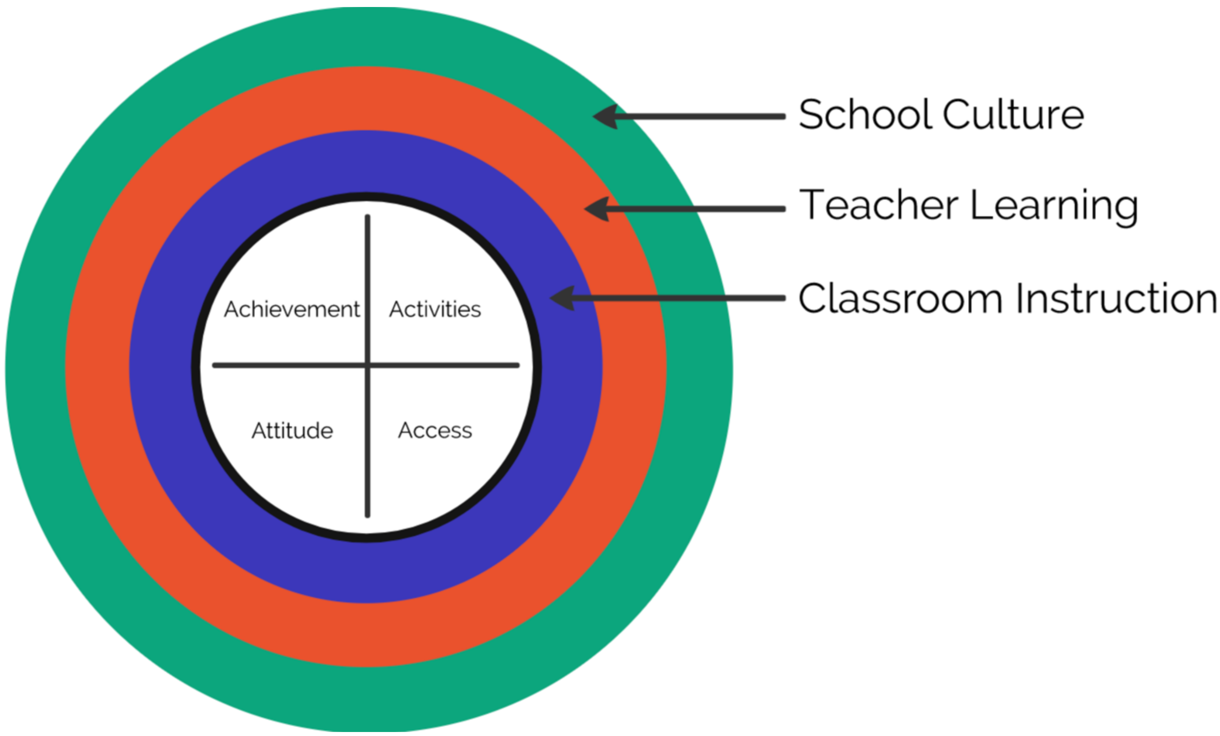


Figure 3.1. *A Nationwide Examination of the Impact of School and Classroom Factors on Middle Grades Students' STEM Pipeline Experiences.*

Method

The 2012 National Survey of Science and Mathematics Education (NSSME, Weis & Banilower, 2014) was designed to examine trends in mathematics and science education across grades K-12 regarding teacher background and experience, curriculum and instruction, and the availability and use of instructional resources. The nationally representative survey used a two-stage stratified probability sample. First, 2,000 schools were sampled within strata and 10,000 teachers were sampled within the selected schools. Due to smaller numbers of teachers in advanced subjects, this group of teachers was oversampled in order to have enough respondents. The NSSME surveyed a total of 7,752 science and mathematics teachers and 1,504 schools across the United States. In addition to the teacher surveys, the NSSME surveyed each school from which teachers were sampled. The NSSME Mathematics and Science Program Questionnaires focused on school-level programmatic information, such as program types, percentage of students taking Algebra I in eighth grade, and school programs and policies.

Horizon Research granted permission for use of the public use NSSME data. All datasets were downloaded as SPSS version 23 files. The data included four datasets and related questionnaires, including Mathematics and Science Teacher datasets and Questionnaires as well as the Mathematics and Science Program datasets and Questionnaires. The Teacher Questionnaires for both mathematics and science included five sections: (a) teacher background and opinions on various instructional topics, (b) teachers' views on their mathematics or science instruction, (c) teachers' reflections on their most recently completed unit, (d) teachers' reflections on their most recently completed lesson, and (e) demographic information. The Program and Teacher surveys were merged using the school identification number (NSSCHLID) for both mathematics and science. Due to slightly different questions on each survey, mathematics and science were separately analyzed.

The study's sample included a sub-sample of full-time public-school mathematics and science teachers from across the U.S. teaching in grades 5 through 8 from the full NSSME sample. Due to the survey's use of two forms (Matrix A and Matrix B) with non-overlapping items, the sub-sample was chosen from only teachers who completed the Matrix A version of the survey, as the majority of the items of interest were asked on the Matrix A survey form. The study's final sample included 2,778 teachers from 1,162 campuses. The sample was just over 50% mathematics teachers (mathematics n = 1,447/ 52.1%; science n = 1,331/ 47.9%) and relatively experienced, with the majority of teachers in both subject groups having 15 or more years of teaching experience. Due to the nested nature of teachers within schools in the dataset and based on previous work with similar independent and dependent measures, the study utilized hierarchical linear modeling, where appropriate, to examine the proportion of variance in teachers' perceptions of their instructional objectives, instructional practices, and use of instructional technology. In cases where differences across schools did not explain a significant amount of variance, multiple linear regression was used to examine the relationship of study predictors on target outcomes.

Data Analysis

Due to the nested nature of teachers within schools, the study utilized multi-level modeling to examine the proportion of variance in STEM teacher perceptions across schools. In cases where null models did not indicate that a significant amount of the variance in a dependent measure was explained by differences across schools, multiple linear regression was used to examine the impact of predictors on outcomes. The teacher and program survey files were merged using the school identification number (NSSCHLID) in SPSS 23 for both mathematics and science. Once the files were merged in SPSS, they were imported into Stata 15.0 and survey set with appropriate jackknife teacher or class level replicate weights in order to correctly account for standard errors due to non-random sampling. Stata 15.0 software was used for final hierarchical linear modeling (HLM) due to the fact that HLM cannot be conducted in SPSS. As mentioned previously, mathematics and science were analyzed using separate statistical models due to slightly different questions on each subject-area survey.

Models

Two-level hierarchical linear modeling with random intercepts was utilized for all dependent measures. The cluster identifier was each school's identification number for NSSME administration (NSSCHLID). There were a total of 955 clusters for math and 918 clusters for science. In each model, the slopes of both school- and teacher-level predictors were held constant, while the intercepts were allowed to vary across schools and teachers. Descriptive statistics for all predictors were analyzed and predictors with large amounts of missing data and/or little variation were eliminated from final models. In addition, listwise deletion was used for with missing data on the included measures (Enders, 2010). Raudenbush and Bryk's (2002) R^2 model was used to calculate total variance explained by the two-level models. The teacher-level predictors utilized across all final models included β_{10}*ContentPDType*$_{ij}$ (type of content-based professional development received), β_{11}*STEMDegree*$_{ij}$ (whether teachers had a STEM-related bachelor's degree), β_{12}*K-12Experience*$_{ij}$ (total years of teaching experience), β_{13}*GradeLevelExperience*$_{ij}$ (total years of experience in one's current grade level), β_{14}*TotalPDTime*$_{ij}$ (total number of hours of professional development received in the last three years), β_{15}*EquipSupplies*$_{ij}$ (teachers' perceptions of the adequacy of equipment and supplies on their campus), and β_{16}*TechProblems* (teachers' perceptions of the degree to which technology problems interfere with instruction). Two predictors, β_{18}*TchrLeadership*$_{ij}$ (teachers' opportunities to participate in leadership activities such as mentoring or coaching), and β_{19}*ContentPreparedness*$_{ij}$ (teachers' perceptions of their level of preparedness for various aspects of content) were included in the mathematics models only due to large amounts of missing data in the science teacher sample (β_{19}*ContentPreparedness*$_{i}$) or items not included on the STQ Matrix A (β_{17}*Pdemphasis*$_{ij}$ [reform-based emphasis of PD] and β_{18}*TchrLeadership*$_{ij}$). One item, β_{13}*InstruPreparedness*$_{ij}$ (teachers' perceptions of their feelings of preparedness with reform-based instruction) was included in the science model only as it was not asked on the MTQ Matrix A. Finally, one teacher-level predictor, β_{17}*Pdemphasis*$_{ij}$, was excluded from the final models due to large amounts of missing data.

School-level predictors utilized in each model included γ_{01} *ExternalPartnerships* (the number of types of external partnerships in a school), γ_{12} *PDTimeTypes* (the number of different types of PD time allocation in a school), γ_{13}*TeacherStudyGroups* (the availability of teacher study groups for PD), γ_{14}*ContentSpecificPD* (whether or not a school provided content-specific PD for teachers) and γ_{15}*CoachingAvailability* (whether or not schools had instructional coaching available to teachers). Two school-level predictors, *InstruBudget* (total annual content budget) and *CoachingByTeachers* (whether a school had teachers participating in instructional coaching), were excluded from the final models in both mathematics and science prior to analysis due to large amounts of missing data. For each model, both variation across clusters (U_{0j}), and variation within schools (e_{ij}) were included. Table 3.1 shows the null and final models used in the study's analysis for mathematics and science.

Table 3.1

Two-level Hierarchical Linear Models for Dependent Measures

Dependent Measure	Model
Instructional Objectives	Null Model $InstruObjectives_{ij} = \gamma_{00} + U_{0j} + e_{ij}$ Two-level Model (Mathematics): $InstruObjectives_{ij} = \beta_{01} + \beta_{10}ContentPDType_{ij} + \beta_{11}STEMDegree_{ij} + \beta_{12}K\text{-}12Experience_{ij} + \beta_{13}GradeLevelExperience_{ij} + \beta_{14}TotalPDTime_{ij} + \beta_{15}EquipSupplies_{ij} + \beta_{16}TechProblems + \beta_{18}TchrLeadership_{ij} + \beta_{19}ContentPreparedness_{ij} + \gamma_{01}\,ExternalPartnerships + \gamma_{12}PDTimeTypes + \gamma_{13}TeacherStudyGroups + \gamma_{14}ContentSpecificPD + \gamma_{15}CoachingAvailability + U_{0j} + e_{ij}$

Table 3.1 Continued

Dependent Measure	Model
	Two-level Model (Science):
	$InstruObjectives_{ij} = \beta_{01} + \beta_{10}ContentPDType_{ij} + \beta_{11}STEMDegree_{ij} + \beta_{12}K\text{-}12Experience_{ij} + \beta_{13}GradeLevelExperience_{ij} + \beta_{14}TotalPDTime_{ij} + \beta_{15}EquipSupplies_{ij} + \beta_{16}TechProblems + \gamma_{01}ExternalPartnerships + \gamma_{12}PDTimeTypes + \gamma_{13}TeacherStudyGroups + \gamma_{14}ContentSpecificPD + \gamma_{15}CoachingAvailability + U_{0j} + e_{ij}$
Instructional Practices	Null Model
	$InstruPractices_{ij} = \gamma_{00} + U_{0j} + e_{ij}$
	Two-level Model (Mathematics):
	$InstruObjectives_{ij} = \beta_{01} + \beta_{10}ContentPDType_{ij} + \beta_{11}STEMDegree_{ij} + \beta_{12}K\text{-}12Experience_{ij} + \beta_{13}GradeLevelExperience_{ij} + \beta_{14}TotalPDTime_{ij} + \beta_{15}EquipSupplies_{ij} + \beta_{16}TechProblems + \beta_{18}TchrLeadership_{ij} + \beta_{19}ContentPreparedness_{ij} + \gamma_{01}ExternalPartnerships + \gamma_{12}PDTimeTypes + \gamma_{13}TeacherStudyGroups + \gamma_{14}ContentSpecificPD + \gamma_{15}CoachingAvailability + U_{0j} + e_{ij}$
	Two-level Model (Science):
	$InstruObjectives_{ij} = \beta_{01} + \beta_{10}ContentPDType_{ij} + \beta_{11}STEMDegree_{ij} +$

Table 3.1 Continued

Dependent Measure	Model
	β_{12}*K-12Experience*$_{ij}$ + β_{13}*GradeLevelExperience*$_{ij}$ + β_{14}*TotalPDTime*$_{ij}$ + β_{15}*EquipSupplies*$_{ij}$ + β_{16}*TechProblems* + γ_{01} *ExternalPartnerships* + γ_{12}*PDTimeTypes* + γ_{13}*TeacherStudyGroups* + γ_{14}*ContentSpecificPD* + γ_{15}*CoachingAvailability* + U_{0j} + e_{ij}
Instructional Technology Use	Null Model *InstruTechUse*$_{ij}$ = γ_{00} + U_{0j} + e_{ij} Two-level Model (Mathematics): *InstruObjectives*$_{ij}$ = β_{01} + β_{10}*ContentPDType*$_{ij}$ + β_{11}*STEMDegree*$_{ij}$ + β_{12}*K-12Experience*$_{ij}$ + β_{13}*GradeLevelExperience*$_{ij}$ + β_{14}*TotalPDTime*$_{ij}$ + β_{15}*EquipSupplies*$_{ij}$ + β_{16}*TechProblems* + β_{18}*TchrLeadership*$_{ij}$ + β_{19}*ContentPreparedness*$_{ij}$ + γ_{01} *ExternalPartnerships* + γ_{12}*PDTimeTypes* + γ_{13}*TeacherStudyGroups* + γ_{14}*ContentSpecificPD* + γ_{15}*CoachingAvailability* + U_{0j} + e_{ij} Two-level Model (Science): *InstruObjectives*$_{ij}$ = β_{01} + β_{10}*ContentPDType*$_{ij}$ + β_{11}*STEMDegree*$_{ij}$ + β_{12}*K-12Experience*$_{ij}$ + β_{13}*GradeLevelExperience*$_{ij}$ + β_{14}*TotalPDTime*$_{ij}$

Table 3.1 Continued

Dependent Measure	Model
	$+ \beta_{15}$*EquipSupplies*$_{ij}$ $+ \beta_{16}$*TechProblems* $+ \gamma_{01}$ *ExternalPartnerships* $+ \gamma_{12}$*PDTimeTypes* $+ \gamma_{13}$*TeacherStudyGroups* $+ \gamma_{14}$*ContentSpecificPD* $+ \gamma_{15}$*CoachingAvailability* $+ U_{0j} + e_{ij}$

Dependent Measures

The dependent variables focused on three main areas of instruction: (1) teachers' instructional objectives, (2) instructional practices, and (3) use of instructional technology. These three areas have been identified as instrumental in increasing students' interest and achievement in STEM (Brotman & Moore, 2008; Brown, Concannon, Marx, Donaldson, & Black, 2016; Christensen, Knezek, & Tyler-Wood, 2015; Dare & Roehrig, 2016; Hansen & Gonzalez, 2014; Kim, 2016; Nugent et al., 2015; Riegle-Crumb, Moore, & Ramos-Wada, 2010). Dependent measures were determined using items from the NSSME Mathematics Teacher Questionnaire (MTQ) and Science Teacher Questionnaire (STQ). Principal-components factor analyses with Varimax rotation using Stata 15 statistical analysis software were conducted on items from the questionnaire addressing teachers' perceptions of their: (1) instructional objectives (MTQ items 36a – h; STQ items 49a – g), (2) instructional practices (MTQ items 37a – p; STQ items 50a – 0), and (3) instructional technology use (MTQ items 40a – h; STQ items 53a – g). Eigenvalues greater than 1.00 were used to determine the number of factors, with each item's highest factor loading determining its scale. A regression-based factor score was predicted from the items on each scale. Cronbach's alpha was calculated to determine the internal consistency of the items composing each scale. The results of the factor analysis, factor loadings, and scale reliabilities for mathematics and science are shown in Tables 2 – 7. Factors with Cronbach's alpha reliabilities of less than 0.65 were dropped from analysis due to low reliabilities (Loewenthal, 2001).

Mathematics Instructional Objectives. Mathematics survey respondents were asked to indicate how much emphasis they placed on eight instructional objectives over the course of the school year. All items were scored on a 4-point Likert-type measure with 1 = None, 2 = Minimal emphasis, 3 = Moderate emphasis, and 4 = Heavy emphasis. The analysis yielded two factors with Eigenvalues above 1.00, accounting for 54.0% of the variance. Item factor loadings ranged from 0.123 to 0.770. Factor 1 was labeled Mathematics Reform-Based Objectives (MRBO), focusing mainly on conceptual understanding and increasing student understanding in mathematics, while Factor 2, labeled Mathematics Procedurally-focused Objectives (MPFO), focused mainly on procedural understanding and test preparation. In addition, the internal consistency reliability (Cronbach alpha) of each scale was calculated. The internal consistency reliability coefficients of the scales were 0.59 to 0.72, with an acceptable alpha level for the RBO factor only. The PFO factor was not retained for analysis due to its low reliability. Table 3.2 exhibits the items and their corresponding factor loadings.

Table 3.2

Mathematics Instructional Objectives Items and Factor Loadings

	Factor Loadings	
	MRBO	MPFO
Understanding mathematical ideas	0.625	
Learning mathematical practices	0.745	
Learning about real-life applications of mathematics	0.761	
Increasing students' interest in mathematics	0.722	
Preparing for further study in mathematics	0.451	
Learning mathematical procedures and/or algorithms		0.770
Learning to perform computations with speed and accuracy		0.754
Learning test-taking skills/strategies		0.123
Eigenvalue	2.75	1.03
Cronbach's alpha reliability	0.72	0.59
Total variance explained by factors	54.0%	

Science Instructional Objectives. Science survey respondents were asked to indicate how much emphasis they placed on seven instructional objectives over the course of the school year. The science items were scored on a 4-point Likert-type measure with 1 = None, 2 = Minimal emphasis, 3 = Moderate emphasis, and 4 = Heavy emphasis. The analysis yielded two factors with Eigenvalues above 1.00, accounting for 50.7% of the variance. Item factor loadings ranged from 0.560 to 0.835. Factor 1, similar to mathematics, was labeled Science Reform-Based Objectives (SRBO), as it also focused mainly on conceptual understanding, science processes, and increasing student understanding. Also similar to the mathematics instructional objectives, Factor 2 was labeled Procedurally-focused Objectives (SPFO), as it, too, focused mainly on memorization and test preparation. In addition, the internal consistency reliability (Cronbach alpha) of each scale was calculated. The internal consistency reliability coefficients of the scales were 0.45 to 0.68. The SRBO factor was retained due to its moderate reliability. The SPFO factor was not retained due to low reliability. Table 3.3 shows the items and their corresponding factor loadings.

Table 3.3

Science Instructional Objectives Items and Factor Loadings

	Factor Loadings	
	SRBO	SPFO
Understanding science concepts	0.560	
Learning science process skills	0.669	
Learning about real-life applications of science	0.682	
Increasing students' interest in science	0.740	

Table 3.3 Continued

	Factor Loadings	
	SRBO	SPFO
Preparing for further study in science	0.618	
Memorizing science vocabulary and/or facts		0.835
Learning test taking skills/strategies		0.750
Eigenvalue	2.36	1.19
Cronbach's alpha reliability	0.68	0.45
Total variance explained by factors	50.7%	

Mathematics Instructional Practices. Mathematics survey respondents were asked to indicate the frequency with which they focused on 16 instructional practices, including items ranging from *engaging the whole class in discussions* to *having students develop mathematical proofs*. All items were scored on a 5-point Likert-type measure with 1 = Never, 2 = Rarely (a few times a year), 3 = Sometimes (once or twice a month), 4 = Often (once or twice a week), and 5 = All or almost all mathematics lessons. The analysis yielded four factors with Eigenvalues above 1.00, accounting for 50.5% of the variance. Item factor loadings ranged from 0.414 to 0.794. Factor 1 was labeled Mathematics Reform-Based Practices (MRBP), focusing mainly on practices related to building conceptual understanding and reasoning. Factor 2, labeled Mathematics Reading and Test Preparation (MRTP), focused on reading from a textbook and test preparation. Factors 3 (Mathematics Whole Class Instruction [MWCI]) and 4

(Mathematics Assessment [MA]) focused on whole group instruction and assessment. The internal consistency reliability coefficients of the scales ranged from 0.49 to 0.79, with an acceptable alpha level for the MRBP factor only. The other three factors were not retained for analysis due their low reliabilities. Table 3.4 exhibits the items and their corresponding factor loadings.

Table 3.4

Mathematics Instructional Practices Items and Factor Loadings

	Factor Loadings			
	MRBP	MRTP	MWCI	MA
Have students consider multiple representations in solving a problem	0.551			
Have students explain and justify their method for solving a problem	0.645			
Have students compare and contrast different methods for solving a problem	0.623			
Have students develop mathematical proofs	0.438			
Have students present their solution strategies to the rest of the class	0.638			
Have students write their reflections class or for homework	0.552			

Table 3.4 Continued

	Factor Loadings			
	MRBP	MRTP	MWCI	MA
Have students work in small groups	0.646			
Provide manipulatives for students to use in problem-solving/investigations	0.651			
Have students read from a mathematics textbook/program or other mathematics-related material in class, either aloud or to themselves		0.437		
Give tests and/or quizzes that are predominantly short-answer		0.756		
Focus on literacy skills		0.456		
Have students practice for standardized tests		0.703		
Have students attend presentations by guest speakers focused on mathematics in the workplace		0.414		
Explain mathematical ideas to the whole class			0.797	
Engage the whole class in discussions			0.645	

Table 3.4 Continued

	Factor Loadings			
	MRBP	MRTP	MWCI	MA
Give tests and/or quizzes that include constructed-response/open-ended items				0.794
Eigenvalue	4.13	1.59	1.29	1.07
Cronbach's alpha reliability	0.79	0.59	0.49	
Total variance explained by factors	50.5%			

Science Instructional Practices. Science teacher survey respondents were asked to indicate the frequency with which they focused on 15 instructional practices, many of which were similar to the mathematics instructional practices. All items were scored on the same 5-point Likert-type measure with 1 = Never, 2 = Rarely (a few times a year), 3 = Sometimes (once or twice a month), 4 = Often (once or twice a week), and 5 = All or almost all science lessons. The analysis yielded five factors with Eigenvalues above 1.00, accounting for 59.5% of the variance. Item factor loadings ranged from 0.497 to 0.799. Whereas the mathematics items factored more clearly into reform-based and non-reform based instructional practices, the science reform-based practices split into several factors. Factor 1 was labeled Science Student-focused Practices (SSFP), focused mainly on practices related to student-focused lab experiences and group work, while Factor 2, labeled Formal Inquiry Science Practices (FISP), focused more on formal strategies, such as project-based learning and formal presentations by students or guests. Factors 3 (Science Reading and Test Preparation, SRTP) 4 (Science Assessment, SA), and 5 (Whole Group Instruction, WGI) focused on a range of teacher-directed practices. The internal consistency reliability coefficients of the scales ranged from 0.48 to 0.69, with an acceptable alpha level for the SSFP and FISP factors only. The other three factors were not retained for analysis due their low reliabilities. Table 3.5 exhibits the items and their corresponding factor loadings.

Table 3.5

Science Instructional Practices Items and Factor Loadings

	Factor Loadings				
	SSFP	FISP	SRTP	SA	WGI
Have students work in small groups	0.786				
Do hands-on/laboratory activities	0.731				
Have students represent and/or analyze data using tables, charts, or graphs	0.575				
Require students to supply evidence in support of their claims	0.583				
Engage the class in project-based learning (PBL) activities		0.691			

Table 3.5 Continued

	Factor Loadings				
	SSFP	FISP	SRTP	SA	WGI
Have students make formal presentations to the rest of the class		0.768			
Have students attend presentations by guest speakers focused on science and/or engineering in the workplace		0.762			
Have students read from a science textbook, module, or other science-related material in class, either aloud or to themselves			0.799		
Have students write their reflections in class or for homework			0.563		
Focus on literacy skills			0.629		
Give tests and/or quizzes that include constructed-response/open-ended items				0.497	

Table 3.5 Continued

	Factor Loadings				
	SSFP	FISP	SRTP	SA	WGI
Give tests and/or quizzes that are predominantly short-answer				0.672	
Have students practice for standardized tests				0.603	
Explain science ideas to the whole class					0.789
Engage the whole class in discussions					0.826
Eigenvalue	3.60	1.86	1.36	1.00	1.03
Cronbach's alpha reliability	0.68	0.69	0.62	0.48	0.50
Total variance explained by factors	59.5%				

Mathematics Instructional Technology Use. The final dependent measure focused on the frequency with which the teachers surveyed used different types of instructional technology. Survey respondents were asked to indicate the frequency with which they used eight types of technology, ranging from laptop computers to classroom response devices, or "clickers." All items were scored on a 5-point Likert-type measure with 1 = Never, 2 = Rarely (a few times a year), 3 = Sometimes (once or twice a month), 4 = Often (once or twice a week), and 5 = All or almost all mathematics lessons. The analysis yielded three factors with Eigenvalues above 1.00, accounting for 60.0% of the variance. Item factor loadings ranged from 0.379 to 0.900. Factor 1, labeled Routine Mathematics Technology (RMT), contained three items, personal computers, hand-held computers, and internet. Factor 2 (CALC) included four function calculators and graphing calculators and Factor 3 (ACTV) included graphing calculators, data collection probes, and classroom response systems. The internal consistency reliability coefficients of the scales ranged from 0.15 to 0.73, with an acceptable alpha level for the RMT factor only. The other factors were not retained for analysis due their low reliabilities. Table 3.6 exhibits the items and their corresponding factor loadings.

Table 3.6

Mathematics Instructional Technology Use Items and Factor Loadings

	Factor Loadings		
	RMT	CALC	ACTV
Personal computers, including laptops	0.900		
Hand-held computers	0.479		
Internet	0.899		
Four-function calculators		0.812	
Scientific calculators		0.789	
Graphing calculators			0.692
Probes for collecting data			0.726
Classroom response system, or “Clickers”			0.379

Table 3.6 Continued

Eigenvalue	2.19	1.41	1.20
Cronbach's alpha reliability	0.73	0.54	0.15
Total variance explained by factors	60.0%		

Science Instructional Technology Use. The science instructional technology use dependent measure also focused on the frequency with which the teachers surveyed used different types of instructional technology. Survey respondents were asked to indicate the frequency with which they used five types of technology, including laptop computers and classroom response devices, or "clickers." One item, calculators, was omitted because it was only provided to K-5 teachers. All items were scored on a 5-point Likert-type measure with 1 = Never, 2 = Rarely (a few times a year), 3 = Sometimes (once or twice a month), 4 = Often (once or twice a week), and 5 = All or almost all mathematics lessons. The analysis yielded two factors with Eigenvalues above 1.00, accounting for 63.4% of the variance. Item factor loadings ranged from 0.519 to 0.896. Factor 1, labeled Routine Science Technology (RST), contained the same three items and the mathematics instructional technology use initial factor: personal computers, hand-held computers, and internet. Factor 3 (ACTV) included data collection probes and classroom response systems. The internal consistency reliability coefficients of the two scales were 0.72 and 0.33, respectively. Only the RST scale was retained for analysis. Table 3.7 shows the items and their corresponding factor loadings.

Table 3.7

Science Instructional Technology Use Items and Factor Loadings

	Factor Loadings	
	RST	ACTV
Personal computers, including laptops	0.892	
Hand-held computers	0.519	
Internet	0.896	
Calculators		0.701
Probes for collecting data		0.786
Eigenvalue	2.17	1.01
Cronbach's alpha reliability	0.72	0.33
Total variance explained by factors	63.4%	

Independent Measures

Due to the study's overarching purpose to identify alterable and actionable factors as called for by leading federal STEM research agencies (Institute for Education Sciences and the National Science Foundation, 2013; National Academies of Sciences, Engineering, and Medicine, 2015), the study's predictors focused on malleable school- and teacher-level factors that have been identified as impactful in both general and STEM-specific changes in teachers' instructional practices and subsequent improvements in student outcomes. The predictors are mainly derived from the *Contexts for Teachers' Learning* framework (National Academies of Sciences, Engineering, and

Medicine's Committee on Strengthening Science Education, 2015). The framework is based on a synthesis of over a decade of work in school- and teacher-level factors that have been shown to create supportive contexts for STEM teaching and learning. The framework is based on the work of Bryk and colleagues at the Chicago Consortium for School Reform (Bryk, Sebring, Allensworth, Suppescu, & Easton, 2010), as well as multiple studies of comprehensive school reform efforts (Cohen & Hill, 2000, 2001; Cohen, Raudenbush, & Ball, 2003; Le et al, 2006; Rowan, Corenti, Miller, & Camburn, 2009).

The four main aspects of the framework are: (1) professional capacity, (2) coherent instructional guidance, (3) leadership, and (4) time and funding. School-level predictors addressing each of the four areas of the framework were selected from the NSSME mathematics or science program questionnaires, while teacher-level predictors were selected from the NSSME mathematics or science teacher questionnaires. Categorical survey items were analyzed as categorical data, while several numeric items, such as years of teaching experience, were banded into categories to examine commonalities within sub-groups of teachers. The remaining predictors were composite variables of sub-items addressing similar constructs. For ease of interpretation, all composite variables were scaled by summing values across all items in a scale and dividing by the total possible value for all items, assigning each composite variable a value between zero and 1. Below is a brief description of each set of predictors.

Professional capacity. Professional capacity focuses on efforts by teachers and schools to build the instructional and collaborative capacity of staff through a variety of means, including collaboration, staff qualifications, and partnerships. Seven predictors addressed professional capacity, three of which were school-level and five of which were teacher-level. The school-level predictors included the number of different types of external partnerships a school engaged in (ExternalPartnerships), ranging from family math/science nights to bringing in outside STEM professionals or mentors, and the availability of teacher study groups (TeacherStudyGroups). The five school-level predictors included whether or not a teacher had a STEM degree (STEMDegree), years of teaching experience in general (TotalExperience) and in their particular grade level (GradeLevelExperience), teachers' perceptions of their level of content preparedness (ContentPreparedness), and teachers' perceptions of their level of instructional preparedness (InstrPreparedness).

Coherent instructional guidance. Coherent instructional guidance focuses on the degree to which teachers have opportunities to learn new practices and consider ways in which new practices might be adapted for successful implementation in their classroom and school contexts. Four predictors addressed coherent instructional guidance. The two school-level predictors focused on whether or not a school provided opportunities for teachers to engage in content-specific professional development (ContentSpecificPD) and work with instructional coaches (CoachingAvailability). The two teacher-level predictors included the whether or not a teacher had participated in content-specific professional development in the last three years ([ContentPDType], i.e., workshop or conference, professional association meeting, professional learning community) and the degree to which the professional development aligned with research-based aspects of high-quality STEM professional development ([PDEmphasis], Garet et al., 2010; Loucks-Horsley, Stiles, Mundry, Love, & Hewson 2010; Wilson 2011), including opportunities to engage in investigations, examine student artifacts, and reflect on strategies after trying them in their classrooms.

Leadership. The leadership aspect of the *Contexts for Teachers' Learning* framework heavily emphasizes the importance of principal leadership for school improvement, with a secondary focus on teacher leadership. The NSSME dataset did not examine principal leadership in sufficient detail, therefore, predictors of principal leadership were not included in the present study. The two predictors, instead, focused on teacher leadership, an important and under-researched area of leadership in STEM. The school-level predictor (CoachingByTeachers) focused on the extent to which teachers provided instructional coaching, either full- or part-time, on a campus. The teacher-level predictor (TchrLeadership) focused on the whether or not a teacher had participated in a teacher leadership role in the last three years.

Time and funding. Time and funding may be viewed as non-malleable and non-alterable in many cases. One could argue, however, that proper allocation of both of these factors contribute heavily to the success or failure of improvement in students' STEM outcomes. Though schools have little control over allocation of total amounts of time and funding, an examination of these areas of concern could result in potential re-allocation of both of these resources within individual campuses and/or help schools and school systems better understand how, or if, time and funding are impacting teachers' instructional practices. The two school-level predictors included a school's total annual

instructional budget (InstrBudget) and whether or not schools utilized one or more types of time ([PDTimeTypes], e.g., early release for students, common planning times, etc.) to create opportunities for content-focused professional development. The three teacher-level predictors focused on teachers' total amount of professional development time in the last three years (TotalPDTime), teachers' perceptions of the availability of equipment and supplies (EquipSupplies), and teachers' perceptions of the degree to which instructional technology was *not* a barrier to instruction (TechProblems).
Table 3.8 provides information regarding the data type, NSSME items utilized, and descriptive statistics of each of the predictors used. The mean and standard deviation is given for all scaled items and the number and percent of responses in each category is provided for categorical and dichotomous items.

Table 3.8

School- and Teacher-Level Predictors – Data Types, NSSME Items, and Sample Descriptive Statistics

	Data Type	NSSME Item(s)	Mathematics		Science	
Professional Capacity			M/n	SD/%	M/n	SD/%
School-level predictors						
ExternalPartnerships	Scale	MPQ5a-i/SPQ5a-k	0.38	0.20	0.48	0.24
TeacherStudyGroups	Dichotomous	MPQ27/SPQ41				
Yes			793	60.81	715	59.34
No			511	39.19	490	40.66
Teacher-level predictors						
STEMDegree	Dichotomous	MTQ14b-e/MTQ14b				
Yes			728	50.31	708	53.19
No			719	49.69	623	46.81
TotalExperience	Categorical	MTQ1a/STQ1a				
Less than 4 years			257	19.68	251	21.05
5-9 years			283	21.67	281	23.01

Table 3.8 Continued

	Data Type	NSSME Item(s)	Mathematics		Science	
10-14 years			247	18.91	244	19.98
15 or more years			519	39.74	439	39.95
GradeLevelExperience	Categorical	MTQ1b/STQ1b				
Less than 4 years			292	20.45	311	23.70
5-9 years			343	24.02	309	23.55
10-14 years			272	19.05	269	20.50
15 or more years			521	36.48	423	32.24
ContentPreparedness	Scale	MTQ27a-h/STQ40a-f	0.88	0.11	*	*
InstrPreparedness	Scale	MTQ28a-j/STQ41a-j	*	*	0.77	0.14
Coherent Instructional Guidance			M/n	SD/%	M/n	SD/%
School-level predictors						
ContentSpecificPD	Dichotomous	MPQ25/SPQ39				
Yes			864	66.11	646	53.57
No			443	33.89	560	46.43

Table 3.8 Continued

	Data Type	NSSME Item(s)	Mathematics		Science	
CoachingAvailability	Dichotomous	MPQ40/SPQ54				
Yes			436	33.49	362	30.04
No			866	66.51	843	69.96
Teacher-level predictors						
ContentPDType	Dichotomous	MTQ20a-c/STQ20a-c				
Yes			1,297	89.63	1,143	85.88
No			150	10.37	188	14.12
PDEmphasis	Scale	MTQ22a-f/STQ35a-f	0.64	0.13	*	*
Leadership			M/n	SD/%	M/n	SD/%
School-level predictors						
CoachingByTeachers	Scale	MTQ44a-f/STQ58a-f	0.47	0.17	0.45	0.14
Teacher-level predictors						
TchrLeadership	Dichotomous	MTQ26b-e/STQ38b-e				
Yes			791	54.66	*	*

Table 3.8 Continued

	Data Type	NSSME Item(s)	Mathematics		Science	
No			656	45.34	*	*
Time and Funding			M/n	SD/%	M/n	SD/%
School-level predictors						
InstrBudget	Scale	MPQ19a-c/SPQ31a-c	$4,039.63	$7,437.56	$7,955.08	$11,156.30
PDTimeTypes	Dichotomous	MPQ39a-f/SPQ53a-f				
Yes			1,304	90.12	1,201	90.23
No			143	9.88	130	9.77
Teacher-level predictors						
TotalPDTime	Categorical	MTQ21/STQ34				
No hours			136	9.41	153	11.53
Less than 6 hours			115	7.96	111	8.36
6-15 hours			319	22.08	298	22.46
16-35 hours			344	23.81	273	20.57

Table 3.8 Continued

	Data Type	NSSME Item(s)	Mathematics		Science	
More than 35 hours			531	36.75	492	37.08
EquipSupplies	Scale	MTQ49a-d/STQ63-66	0.76	0.19	0.71	0.20
TechProblems	Scale	MTQ50a-f/STQ67a-g	0.87	0.15	0.83	0.16

Source. National Survey of Science and Mathematics Educators Teacher and Program Surveys 2012–13.
Note. "MPQ" = Mathematics Program Questionnaire, "SPQ" = Science Program Questionnaire, "MTQ" = Mathematics Teacher Questionnaire, "STQ" = Science Teacher Questionnaire, "*" indicates items only on Matrix A MTA or Matrix A STQ.

Results

In order to determine the extent to which teacher perceptions of their reform-based mathematics or science instructional objectives, instructional practices, and instructional technology use were attributable to differences across schools, intra-class correlation coefficients (ICCs) were calculated for null models for each dependent measure. The ICCs of the dependent measures ranged from 2.69% - 31.50%. Difference across schools explained non-statistically significant amounts of the differences in mathematics and science teachers' perceptions of their use of reform-based instructional objectives (MRBO, ICC = 2.69%, $p = 0.30$, SRBO, ICC = 4.36%, $p = 0.20$). With regard to teachers' use of instructional practices, differences across schools explained statistically significant amounts of the variance in mathematics teacher perceptions of their use of mathematics reform-based practices (MRBP, ICC = 18.77%, $p < .001$), as well as statistically significant amounts of the variance in science teachers' perceptions of student-focused practices (SSFP, ICC = 27.88%, $p < .001$). However, differences across schools explained a nonsignificant amount of the variance in teachers' use of formal inquiry science practices (FISP- 7.69%, $p = 0.09$). Finally, differences across schools explained statistically significant amounts of the variance in both mathematics and science teachers' perceptions of their instructional technology use (mathematics, ICC = 30.86%, $p < .001$, science, ICC = 31.50%, $p < .001$). Therefore, multiple linear regression, rather than HLM, was used to examine the influence of the study's predictors on mathematics and science teachers' reform-based instructional objectives (MRBO and SRBO), and science teachers' use of formal inquiry science practices (FISP).

Mathematics Instructional Objectives

Due to the nonsignificant amount of variance explained by differences across schools on teachers' use of mathematics instructional objectives, multiple linear regression was used to examine the effect of school- and teacher-level factors on teachers' use of effective STEM instructional practices. The total R^2 of the model was .1013, indicating that approximately 10% of the variance in teachers' perceptions of their use of reform-based instructional objectives can be explained by the model's predictors. Three predictors in the areas of Coherent Instructional Guidance, Professional Capacity, and Time and Funding significantly predicted teachers' use of reform-based mathematics instructional objectives. Teachers who had participated in any of the three types of content-specific professional development in the three years prior to taking the survey (*ContentPDType*) had a mathematics reform-based objectives score that was 0.93 points higher on average than teachers who had not participated in content-specific professional development in the last three years. In addition, as teachers' perceptions of their level of content preparedness (*ContentPreparedness*) increased, the use of mathematics reform-based objectives increased by 1.22 points on average. Finally, teachers with less than six hours of content-specific professional development in the last three years (*TotalPDTime*) had a mathematics reform-based objectives score that was 1.01 points lower on average than teachers with no content-specific professional development in the last three years. There were no other significant differences in teachers' perceptions of their use of mathematics reform-based objectives. Table 3.9 provides the multiple regression results for teachers' perceptions of mathematics reform-based objectives.

Table 3.9

Results of Multiple Regression Analysis – Mathematics Reform-based Objectives

Predictor	*β*	SE	*t*	*p*	*95% CI*
Teacher-level Predictors					
ContentPDType	0.929	0.278	3.34	0.001**	0.347 – 1.48
STEMDegree	-0.038	0.085	-0.45	0.655	-0.206 – 0.132
K-12Experience					
5 – 9 years	-0.287	1.27	-0.02	0.982	-2.56 – 2.50
10 – 14 years	0.179	1.05	0.17	0.866	-1.93 – 2.28
15 or more years	0.255	1.21	0.21	0.833	-2.15 – 2.66
GradeLevelExperience					
5 – 9 years	0.072	1.06	0.07	0.946	-2.04 – 2.19
10 – 14 years	-0.073	0.740	-0.11	0.921	-1.55 – 1.40

Table 3.9 Continued

Predictor	β	SE	t	p	*95% CI*
15 or more years	-0.112	0.988	-0.11	0.910	-2.08 – 1.86
TotalPDTime					
Less than 6 hours	-1.013	0.347	-2.92	0.005*	-1.70 – -0.320
6 – 15 hours	-0.683	0.447	-1.53	0.132	-1.58 – 0.211
16 – 35 hours	-0.897	0.485	-1.85	0.069	-1.87 – 0.072
More than 35 hours	-0.651	0.489	-1.33	0.187	-1.63 – 0.324
EquipSupplies	0.792	0.456	1.740	0.087	-0.118 – 1.703
TechProblems	-0.487	0.336	-1.450	0.151	-1.157 – 0.183
TchrLeadership	0.087	0.257	0.34	0.738	-0.427 – 0.600
ContentPreparedness	1.22	0.415	2.94	0.005*	0.391 – 2.05
School-level predictors					
ExternalPartnerships	0.256	1.083	0.240	0.814	-1.908 – 2.419
TeacherStudyGroups	0.013	0.222	0.060	0.955	-0.431 – 0.456

Table 3.9 Continued

Predictor	*β*	SE	*t*	*p*	*95% CI*
ContentSpecificPD	0.003	0.162	0.020	0.984	-0.320 – 0.327
CoachingAvailability	-0.006	0.229	-0.030	0.978	-0.464 – 0.451

Note. $* p < 0.05$, $** p < 0.01$

Science Instructional Objectives

Similar to the mathematics data, a nonsignificant amount of variance explained by differences across schools on teachers' use of science instructional objectives indicated that multiple linear regression should be used to examine study predictors. The total R^2 of the model was .1653, indicating that approximately 17% of the variance in teachers' perceptions of their use of reform-based instructional objectives can be explained by the model's predictors. Only one predictor related to Professional Capacity significantly predicted teachers' use of reform-based science instructional objectives. As teachers' perceptions of their level of instructional preparedness (*InstrPreparedness*) increased, their use of science reform-based objectives increased by 2.73 points on average. There were no other significant differences in science teachers' perceptions of their use of reform-based objectives. Table 3.10 provides the multiple regression results for teachers' perceptions of science reform-based objectives.

Table 3.10

Results of Multiple Regression Analysis – Science Reform-based Objectives

Predictor	β	SE	*t*	*p*	*95% CI*
Teacher-level Predictors					
ContentPDType	0.308	0.283	1.09	0.278	-0.254 – 0.872
STEMDegree	-0.072	0.098	-0.73	0.466	-0.266 – 0.123
K-12Experience					
5 – 9 years	-0.040	0.204	-0.20	0.844	-2.56 – 2.50
10 – 14 years	-0.176	0.225	-0.780	0.437	-0.623 – 0.272
15 or more years	-0.360	0.381	-0.950	0.347	-1.118 – 0.398
GradeLevelExperience					
5 – 9 years	-0.032	0.230	-0.140	0.889	-0.491 – 0.426
10 – 14 years	0.018	0.239	0.070	0.942	-0.458 – 0.493
15 or more years	0.343	0.624	0.550	0.584	-0.899 – 1.586
TotalPDTime					
Less than 6 hours	-0.059	0.338	-0.170	0.862	-0.732 – 0.614

Table 3.10 Continued

Predictor	*β*	SE	*t*	*p*	*95% CI*
6 – 15 hours	-0.026	0.675	-0.040	0.969	-1.371 – 1.319
16 – 35 hours	-0.143	0.414	-0.340	0.731	-0.968 – 0.683
More than 35 hours	-0.069	0.433	-0.160	0.873	-0.932 – 0.793
EquipSupplies	0.115	0.755	0.150	0.879	-1.389 – 1.620
TechProblems	0.038	1.623	0.020	0.981	-3.195 – 3.272
InstrPreparedness	2.729	0.444	6.140	0.001**	1.844 – 3.614
School-level predictors					
ExternalPartnerships	-0.082	0.381	-0.210	0.831	-0.841 – 0.678
TeacherStudyGroups	0.052	0.108	0.480	0.630	-0.163 – 0.268
ContentSpecificPD	-0.053	0.132	-0.400	0.690	-0.315 – 0.210
CoachingAvailability	0.025	0.103	0.240	0.809	-0.181 – 0.231
PDTimeTypes	-0.365	1.376	-0.270	0.792	-3.107 – 2.377

Note. ** $p < 0.01$

Mathematics Instructional Practices

The null model for teachers' perceptions of their use of mathematics reform-based instructional practices (MRBP) was statistically significant, indicating that hierarchical linear modeling was the appropriate statistical method for examining the relationship between the school- and teacher-level predictors and teachers' perceptions of their use of reform-based mathematics instructional practices. The random intercept model for teachers' perceptions of their use of reform-based practices with school- and teacher-level predictors was statistically significantly better than the null model ($\chi^2 = 12.95$, $p < .001$), indicating the school- and teacher-level predictors explained a significant portion of the variance in teacher perceptions of MRBP. The school- and teacher-level predictors explained 13.78% of the variance in teachers' perceptions of MRBP. With regard to school-level predictors, teachers in schools offering instructional coaching (*CoachingAvailability*) had an MRBP score of 0.21 points higher on average than teachers in schools where no instructional coaching was offered. There were no other significant differences in school-level predictors. One school-level predictor, *PDTimeTypes*, was omitted from the model due to collinearity.

Five teacher-level predictors had statistically significant influences on mathematics teachers' use of reform-based instructional practices. Teachers participating in teacher leadership activities (*TeacherLeadership*), such as instructional coaching or mentorship, had an MRBP score that was 0.30 points higher on average than teachers who had not participated in leadership activities. In addition, as teachers' levels of content preparedness increased (*ContentPreparedness*), their use of reform-based instructional practices increased by 0.70 points on average. In contrast, teacher experience appeared to negatively impact teachers' use of reform-based instructional practices, with teachers with 10-14 and more than 15 years of experience in their current grade level (*GradeLevelExperience*) having lower perceptions of their use of reform-based practices, on average, than teachers with less experience. Teachers with 10-14 years of teaching experience in their grade level had an MRBP score that was 0.47 points lower on average than teachers with less than five years of experience, while teachers with 15 or more years of experience had an MRBP score that was 0.49 points lower on average than teachers with less than five years of experience. As teachers' perceptions of the adequacy of their equipment and supplies increased (*EquipSupplies*), their use of reform-based instructional practices increased by 0.84 points on average. In contrast, as teachers' perceptions of their technology problems decreased (*TechProblems*), their use of reform-based instructional practices also decreased by 1.09 points on average. There were no other significant teacher-level differences in perceptions of their use of mathematics reform-based instructional practices. One teacher-level predictor, *PDTimeTypes*, was omitted from the model due to collinearity. Table 3.11 provides the random intercept model results for teachers' perceptions of their use of reform-based instructional practices.

Table 3.11

Results for Random-Intercept Model – Mathematics Reform-based Instructional Practices

Parameters	*Coefficient*	*S.E.*	*z*	*P*
Fixed effects				
γ_{00}	-.403	0.057	-1.26	0.209
School-level predictors				
γ_{01} *ExternalPartnerships*	0.132	0.160	0.27	0.791
γ_{13}*TeacherStudyGroups*	-0.121	0.066	-1.84	0.066
γ_{14}*ContentSpecificPD*	0.097	0.068	1.42	0.155
γ_{15}*CoachingAvailability*	0.209	0.068	3.06	0.002**
Teacher-level predictors				
β_{11}*STEMDegree*	-0.085	0.062	-1.338	0.167
β_{12}*K-12Experience*				
5 – 9 years	-0.079	0.158	-0.50	0.618
10 – 14 years	0.135	0.179	0.75	0.451

Table 3.11 Continued

Parameters	*Coefficient*	*S.E.*	*z*	*P*
15 or more years	0.088	0.175	0.50	0.615
β_{13}*GradeLevelExperience*				
5 – 9 years	-0.240	0.154	-1.56	0.120
10 – 14 years	-0.465	0.177	-2.63	0.009**
15 or more years	-0.489	0.180	-2.72	0.007**
β_{14}*TotalPDTime*				
Less than 6 hours	-0.201	0.316	-0.64	0.525
6 – 15 hours	0.033	0.323	0.10	0.918
16 – 35 hours	0.021	0.325	0.07	0.948
More than 35 hours	0.191	0.325	0.59	0.556
β_{15}*EquipSupplies*	0.839	0.164	5.12	0.001**
β_{16}*TechProblems*	-1.09	0.208	-5.25	0.001**
β_{18}*TchrLeadership*	0.300	0.062	4.84	0.001**
β_{19}*ContentPreparednes*	0.698	0.297	2.35	0.019*

Table 3.11 Continued

Random effects			95% Confidence Interval	
			Lower	Upper
σ_e^2	0.123	0.051	0.054	0.278
σ_U^2	0.762	0.057	0.658	0.884

* *p < 0.05,* ** *p < 0.01*

Science Instructional Practices

Science Instructional Practices- Factor 1 (SSFP). The null model for teachers' perceptions of their use of student-focused instructional practices (SSFP) was statistically significant, indicating that differences across schools explained a significant amount of the variance in student-focused instructional practices and HLM was appropriate. The random intercept model for teachers' perceptions of their use of student-focused practices with school- and teacher-level predictors was statistically significantly better than the null model ($\chi^2 = 28.27$, $p < .001$), indicating the school- and teacher-level predictors explained a significant portion of the variance in teacher perceptions of SSFP. The school- and teacher-level predictors explained 22.48% of the variance in teachers' perceptions of SSFP. None of the other school-level predictors were statistically significant.

Two teacher-level predictors had statistically significant influences on science teachers' use of student-focused instructional practices. As teachers' levels of instructional preparedness increased (*InstrPreparedness*), their use of student-focused instructional practices increased by 1.15 points on average. In addition, as teachers' perceptions of the adequacy of their equipment and supplies increased (*EquipSupplies*), their use of student-focused instructional practices increased by 0.76 points on average. There were no other significant teacher-level differences in perceptions of their use of student-focused instructional practices. One teacher-level predictor, *PDTimeTypes*, was omitted from the model due to collinearity. Table 3.12 provides the random intercept model results for teachers' perceptions of student-focused practices.

Table 3.12

Results for Random-Intercept Model – Science Instructional Practices Factor 1

Parameters	*Coefficient*	*S.E.*	*z*	*P*
Fixed effects				
γ_{00}	-2.24	0.721	-3.10	0.002**
School-level predictors				
γ_{01} *ExternalPartnerships*	0.050	0.141	0.35	0.726
γ_{12} *PDTimeTypes*	0.366	0.673	0.54	0.587
γ_{13} *TeacherStudyGroups*	0.048	0.069	0.70	0.486

Table 3.12 Continued

Parameters	*Coefficient*	*S.E.*	*z*	*P*
γ_{14}*ContentSpecificPD*	0.053	0.069	0.77	0.439
γ_{15}*CoachingAvailability*	0.077	0.073	1.06	0.289
Teacher-level predictors				
β_{11}*STEMDegree*	-0.041	0.064	-0.63	0.528
β_{12}*K-12Experience*				
5 – 9 years	0.245	0.153	1.60	0.110
10 – 14 years	-0.011	0.190	-0.806	0.955
15 or more years	-0.130	0.200	-0.65	0.517
β_{13}*GradeLevelExperience*				
5 – 9 years	-0.066	0.149	-0.44	0.658
10 – 14 years	0.153	0.185	0.83	0.407
15 or more years	0.144	0.201	0.72	0.474
β_{14}*TotalPDTime*				
Less than 6 hours	-0.287	0.217	-1.32	0.187

Table 3.12 Continued

Parameters	*Coefficient*	*S.E.*	*z*	*P*
6 – 15 hours	-0.121	0.233	-0.52	0.601
16 – 35 hours	-0.115	0.238	-0.48	0.628
More than 35 hours	0.039	0.236	0.17	0.867
β_{15}*EquipSupplies*	0.757	0.172	4.41	0.001**
β_{16}*TechProblems*	0.042	0.213	0.20	0.842
β_{19}*InstrPreparedness*	1.15	0.224	5.12	0.001**
Random effects			95% Confidence Interval	
			Lower	Upper
σ_e^2	0.199	0.059	0.111	0.357
σ_U^2	0.68	0.061	0.578	0.815

* *p* < 0.05, ** *p* < 0.01

Science Instructional Practices- Factor 2 (FISP). The null model for teachers' perceptions of their use of formal inquiry science practices (FISP) was not statistically significant, indicating that multiple linear regression was the appropriate statistical method for examining the relationship between the study's predictors and teachers' perceptions of their use of reform-based science instructional practices. The total R^2 of the model was .0769, indicating that approximately 8% of the variance in teachers' perceptions of their use of inquiry-based instructional practices can be explained by the model's predictors. Only one predictor at the teacher level significantly predicted teachers' use of formal inquiry-based science instructional objectives. As teachers' perceptions of their level of instructional preparedness (InstrPreparedness) increased, their use of formal inquiry science practices, such as project-based learning, increased by 1.29 points on average. There were no other significant differences in science teachers' perceptions of their use of formal inquiry science instructional practices. Table 3.13 provides the multiple regression results for teachers' perceptions of formal inquiry-focused science practices.

Table 3.13

Results of Multiple Regression Analysis – Science Instructional Practices Factor 2

Predictor	β	SE	t	p	*95% CI*
Teacher-level Predictors					
ContentPDType	-0.412	0.382	-1.080	0.284	-1.172 – 0.348
STEMDegree	-0.200	0.105	-1.900	0.061	-0.409 – 0.010

Table 3.13 Continued

Predictor	*β*	SE	*t*	*p*	*95% CI*
K-12Experience					
5 – 9 years	-0.165	0.981	-0.170	0.867	-2.118 – 1.789
10 – 14 years	-0.021	1.462	-0.010	0.989	-2.933 – 2.892
15 or more years	0.273	1.013	0.270	0.788	-1.745 – 2.291
GradeLevelExperience					
5 – 9 years	0.125	0.951	0.130	0.896	-1.769 – 2.019
10 – 14 years	-0.173	1.158	-0.150	0.882	-2.479 – 2.133
15 or more years	-0.338	0.640	-0.530	0.598	-1.612 – 0.936
TotalPDTime					
Less than 6 hours	0.721	0.419	1.720	0.089	-0.114 – 1.556
6 – 15 hours	0.556	0.458	1.210	0.229	-0.357 – 1.470
16 – 35 hours	0.749	0.494	1.520	0.133	-0.234 – 1.732
More than 35 hours	0.663	0.552	1.200	0.234	-0.437 – 1.764

Table 3.13 Continued

Predictor	*β*	SE	*t*	*p*	*95% CI*
EquipSupplies	-0.286	0.716	-0.400	0.691	-1.712 – 1.140
TechProblems	-0.180	1.015	-0.180	0.859	-2.202 – 1.842
InstrPreparedness	1.286	0.403	3.190	0.002**	0.483 – 2.089
School-level predictors					
ExternalPartnerships	0.071	0.235	0.300	0.764	-0.397 – 0.539
TeacherStudyGroups	-0.064	0.311	-0.210	0.838	-0.682 – 0.555
ContentSpecificPD	0.125	0.227	0.550	0.584	-0.327 – 0.577
CoachingAvailability	-0.018	0.118	-0.150	0.878	-0.252 – 0.216

Note. ** $p < 0.01$

Mathematics Instructional Technology Use

The null model for mathematics teachers' perceptions of their use of instructional technology was statistically significant, indicating that hierarchical linear modeling was an appropriate method to model the effect of school- and teacher-level predictors on teachers' use of instructional technology. The random intercept model for teachers' perceptions of the frequency of their use of routine mathematics technology

(RMT; i.e., personal computers, hand-held computers, and internet) with school- and teacher-level predictors was statistically significantly better than the null model ($\chi^2 = 23.32$, $p < .001$), indicating the school- and teacher-level predictors explained a significant portion of the variance in teacher perceptions their frequency of instructional technology use. The school- and teacher-level predictors explained 26.09% of the variance in teachers' perceptions of RMT. There were no significant school-level predictors. One school-level predictor, *PDTimeTypes*, was omitted from the model due to collinearity.

Five teacher-level predictors had statistically significant influences on mathematics teachers' use of instructional technology. Teachers participating in teacher leadership activities (*TeacherLeadership*), such as instructional coaching or mentorship, had an RMT score that was 0.21 points higher on average than teachers who had not participated in leadership activities. Regarding teaching experience, teachers with 10-14 years of total K-12 teaching experience (*K-12Experience*) perceived that they used instructional technology more frequently than teachers with five or less years of experience, with an RMT score of about 0.37 points higher on average. In contrast, teachers with 10-14 years and 15 or more years of experience in their current grade level (*GradeLevelExperience*) had RMT scores that were 0.73 and 0.62 points lower on average than teachers with less than five years of experience. In addition, as teachers' perceptions of the adequacy of their equipment and supplies increased (*EquipSupplies*), the frequency of their use of instructional technology increased by 0.78 points on average. However, as teachers' perceptions of their technology problems decreased

(*TechProblems*), their use of reform-based instructional practices also decreased by 0.43 points on average. There were no other significant teacher-level differences in mathematics teachers' perceptions of the frequency of their use of instructional technology. One teacher-level predictor, *ContentPDType*, was omitted from the model due to collinearity. Table 3.14 provides the random intercept model results for teachers' perceptions of their use of instructional technology.

Table 3.14

Results for Random-Intercept Model – Mathematics Instructional Technology Use

Parameters	*Coefficient*	*S.E.*	*z*	*P*
Fixed effects				
γ_{00}	0.022	0.319	0.07	0.946
School-level predictors				
γ_{01} *ExternalPartnerships*	0.019	0.165	0.120	0.908
γ_{13}*TeacherStudyGroups*	-0.079	0.068	-1.170	0.242
γ_{14}*ContentSpecificPD*	0.052	0.070	0.750	0.454
γ_{15}*CoachingAvailability*	-0.021	0.070	-0.300	0.764

Table 3.14 Continued

Parameters	*Coefficient*	*S.E.*	*z*	*P*
Teacher-level predictors				
β_{11}*STEMDegree*	-0.106	0.061	-1.740	0.083
β_{12}*K-12Experience*				
5 – 9 years	-0.064	0.153	-0.420	0.677
10 – 14 years	0.374	0.175	2.140	0.032*
15 or more years	0.193	0.172	1.120	0.262
β_{13}*GradeLevelExperience*				
5 – 9 years	-0.215	0.150	-1.430	0.152
10 – 14 years	-0.726	0.175	-4.160	0.000**
15 or more years	-0.625	0.178	-3.510	0.000**
β_{14}*TotalPDTime*				
Less than 6 hours	0.488	0.303	1.610	0.107

Table 3.14 Continued

Parameters	*Coefficient*	*S.E.*	*z*	*P*
6 – 15 hours	0.427	0.308	1.390	0.166
16 – 35 hours	0.358	0.310	1.160	0.248
More than 35 hours	0.375	0.309	1.210	0.225
β_{15}*EquipSupplies*	0.784	0.162	4.840	0.000**
β_{16}*TechProblems*	-0.431	0.207	-2.090	0.037*
β_{18}*TchrLeadership*	0.208	0.061	3.410	0.001*
β_{19}*ContentPreparednes*	-0.105	0.294	-0.360	0.722
Random effects			95% Confidence Interval	
			Lower	Upper
σ_e^2	0.238	0.051	0.157	0.363
σ_U^2	0.675	0.051	0.583	0.783

* $p < 0.05$, ** $p < 0.01$

Science Instructional Technology Use

The null model for science teachers' perceptions of their use of instructional technology was statistically significant, indicating that differences across schools explained a significant amount of the variance in science teachers' instructional technology use and HLM was appropriate. The random intercept model for teachers' perceptions of the frequency of their use of routine science technology (RST; i.e., personal computers, hand-held computers, and internet) with school- and teacher-level predictors was statistically significantly better than the null model ($\chi^2 = 20.33$, $p < .001$), indicating the school- and teacher-level predictors explained a significant portion of the variance in teacher perceptions their frequency of instructional technology use. The school- and teacher-level predictors explained 26.42% of the variance in teachers' perceptions of RST. One school-level predictor, ExternalPartnerships significantly explained the frequency of teachers' use of instructional technology. As the number of school external partnerships increased at a school, teachers' perceptions of the frequency of their instructional technology use increased by 0.33 points on average. There were no other significant school- predictors.

Two teacher-level predictors had statistically significant influences on science teachers' use of instructional technology. As teachers' perceptions of the adequacy of their equipment and supplies increased (*EquipSupplies*), the frequency of their use of instructional technology increased by 0.71 points on average. In addition, as teachers' perceptions of their technology problems decreased (*TechProblems*), their use of reform-based instructional practices increased by 0.74 points on average. There were no other significant teacher-level differences in science teachers' perceptions of the frequency of their use of instructional technology. One teacher-level predictor, *ContentPDType*, was omitted from the model due to collinearity. Table 3.15 provides the random intercept model results for teachers' perceptions of their use of instructional technology.

Table 3.15

Results for Random-Intercept Model – Science Instructional Technology Use

Parameters	*Coefficient*	*S.E.*	*z*	*P*
Fixed effects				
γ_{00}	-1.58	0.722	-2.19	0.028*
School-level predictors				
γ_{01} *ExternalPartnerships*	0.329	0.139	2.370	0.018*
γ_{12} *PDTimeTypes*	-0.108	0.676	-0.160	0.874
γ_{13}*TeacherStudyGroups*	-0.014	0.068	-0.200	0.839
γ_{14}*ContentSpecificPD*	-0.060	0.068	-0.880	0.379
γ_{15}*CoachingAvailability*	-0.026	0.072	-0.370	0.715
Teacher-level predictors				
β_{11}*STEMDegree*	-0.062	0.063	-0.980	0.326

Table 3.15 Continued

Parameters	*Coefficient*	*S.E.*	*z*	*P*
β_{12}*K-12Experience*				
5 – 9 years	0.118	0.148	0.800	0.425
10 – 14 years	0.122	0.182	0.670	0.501
15 or more years	-0.013	0.194	-0.070	0.946
β_{13}*GradeLevelExperience*				
5 – 9 years	0.034	0.144	0.240	0.814
10 – 14 years	0.021	0.178	0.120	0.906
15 or more years	0.058	0.195	0.300	0.767
β_{14}*TotalPDTime*				
Less than 6 hours	0.075	0.208	0.360	0.719
6 – 15 hours	0.125	0.226	0.550	0.581
16 – 35 hours	0.321	0.230	1.390	0.163
More than 35 hours	0.131	0.228	0.580	0.564

Table 3.15 Continued

Parameters	*Coefficient*	*S.E.*	*z*	*P*
β_{15}*EquipSupplies*	0.707	0.168	4.210	0.001**
β_{16}*TechProblems*	0.742	0.206	3.610	0.001**
β_{19}*InstrPreparednes*	0.231	0.219	1.050	0.293
Random effects			95% Confidence Interval	
			Lower	Upper
σ_e^2	0.237	0.053	0.152	0.368
σ_U^2	0.659	0.053	0.563	0.771

* $p < 0.05$, ** $p < 0.01$

Summary and Discussion

Due to the fact that over 80% of the fastest growing occupations in the United States are in STEM-related fields (U. S. Bureau of Labor Statistics, Employment Projections Program, 2017), a laser-like focus on K-12 STEM improvement has placed STEM teachers' instructional practices and STEM student achievement under a great deal of scrutiny in recent years. Recent research showing that the middle grades is a time when students' attitudes regarding STEM fields and careers are most subject to change (Aschbacher et al., 2014; Catsambis, 1995; Christensen & Knezek, 2016; Hinojosa et al, 2016; Ing & Nylund-Gibson, 2017; Nugent et al, 2015; Oakes, 1990) highlights a critical need for further study of middle grades STEM teachers' practices. However, a majority of current studies focus on the impact of reform-based teaching practice on students' STEM pipeline experiences or on students' self-reported perceptions of their STEM experiences. In addition, though national STEM advocacy groups, such as the American Statistical Association, the National Academy of Sciences, National Academy of Engineering, and Institute of Medicine, advocate frameworks focused on STEM improvement, few studies have provided actionable information for campus and district-level administrators on what alterable school- or teacher-level factors impact the degree to which STEM teachers utilize reform-based instruction. The present study's major findings are summarized below according to *Contexts for Teachers' Learning* framework.

Professional Capacity

Seven study predictors addressed professional capacity, or efforts by teachers and schools to build the instructional and collaborative capacity of staff through a variety of means, including collaboration, staff qualifications, and partnerships. Three predictors were school-level and five were teacher-level. Five predictors related to building teachers' professional capacity significantly predicted study outcomes. One school level predictor, *ExternalPartnerships*, the number of different types of external partnerships a school engaged in, significantly and positively predicted science teachers' frequency of instructional technology use. Additionally, significant teacher-level predictors included *TotalExperience*, *GradeLevelExperience*, *ContentPreparedness* for mathematics teachers, and *InstrPreparedness* for science teachers. Total years of teaching experience significantly and positively predicted mathematics teachers' frequency of instructional technology use for teachers' with 10-14 years of experience compared to teachers with less than five years of experience, while total years of experience in a grade band significantly and negatively predicted both teachers' use of mathematics reform-based practices for teachers with 10 – 14 years and 15 or more years in their grade level compared to teachers with less than five years of experience and mathematics teachers' frequency of instructional technology use for teachers with 10 – 14 years and 15 or more years of experience in their current grade level.

Finally, mathematics teachers' feelings of preparedness with their subject matter (*ContentPreparedness*) significantly and positively predicted teachers' perceptions of their level their use of both mathematics reform-based objectives and practices, while science teachers' perceptions of their instructional preparedness to use a variety of reform-based instructional practices, such as managing classroom discipline and encouraging participation of racial or ethnic minority students in STEM, significantly and positively predicted science teachers' perceptions of their use of science reform-based objectives, student focused practices, and formal inquiry science practices. Two predictors, the availability of teacher study groups (*TeacherStudyGroups*) and whether or not a teacher had a STEM degree (*STEMDegree*), did not significantly predict any of the dependent measures in mathematics or science.

Coherent Instructional Guidance

Coherent instructional guidance focuses on the degree to which teachers have opportunities to learn new practices and consider ways in which new practices might be adapted for successful implementation in their classroom and school contexts. Four predictors addressed coherent instructional guidance. One school-level predictor, *CoachingAvailability*, focused on whether or not a school provided instructional coaches for teachers, significantly and positively predicted mathematics teachers' use of mathematics reform-based practices. Another school-level predictor (*ContentSpecificPD)*, indicating whether or not schools provided opportunities for teachers to engage in content-specific professional development did not significantly predict any of the study's outcomes. In contrast, a similar teacher-level predictor measuring teachers' participation in multiple types of content-focused professional development (*ContentPDType*) did significantly and positively predict mathematics teachers' use of reform-based instructional objectives. A final predictor focused on coherent instructional guidance, *PDEmphasis*, or the degree to which the professional development attended by teachers was aligned with research-based aspects of high-quality STEM professional development, was dropped from analysis due to missing data for the majority of respondents.

Leadership

The leadership aspect of the *Contexts for Teachers' Learning* framework heavily emphasizes the importance of principal leadership for school improvement, with a secondary focus on teacher leadership. The present study focused solely on teacher leadershipdue to the fact that the NSSME dataset did not examine principal leadership. One teacher-level predictor, *TchrLeadership,* focused on whether or not a teacher had participated in a teacher leadership role in the last three years, positively and significantly predicted mathematics teachers' use of reform-based practices and frequency of instructional technology use. However, the school-level leadership predictor, *CoachingbyTeachers,* focused on the extent to which teachers provided instructional coaching, either full- or part-time, on a campus, did not significantly predict any of the study's outcomes for mathematics or science teachers.

Time and Funding

Time and funding may be viewed as non-malleable and non-alterable in many cases. One could argue, however, that proper allocation of both of these factors contribute heavily to the success or failure of improvement in students' STEM outcomes. Though schools have little control over allocation of total amounts of time and funding, an examination of these areas of concern could result in potential re-allocation of both of these resources within individual campuses and/or help schools and school systems better understand how, or if, time and funding are impacting teachers' instructional practices. One predictor, *TotalPDTime,* that served as an indicator of the overall amount of time teachers had participated in professional development in the three years prior to taking the survey was statistically significant. Mathematics teachers with less than six total hours of professional development had a reform-based objectives score that was statistically significantly lower than the reference group of teachers with no hours of professional development in the last three years. It is possible in the case of total professional development time that the difference was related to the fact that the reference group contained mostly young teachers recently graduated from teacher preparation programs and were more likely to be versed in reform-based practices.

There were two predictors focused on funding that had statistically significant impacts on study outcomes. Teachers' perceptions of the adequacy of equipment and supplies related to their content area significantly and positively predicted mathematics teachers' use of reform-based instructional practices and instructional technology use, as well as science teachers' use of student-focused practices and frequency of instructional technology use. In contrast to teachers' perceptions of equipment and supply adequacy, teachers' perceptions of the degree to which technology issues, such as lack of a strong internet connection, were not a problem on their campus, significantly and negatively predicted mathematics teachers' use of reform-based practices and their frequency of instructional technology use. However, science teachers' perceptions of the degree to which technology issues were not a problem on their campus significantly and positively predicted their frequency of instructional technology use. One predictor, *PDTimeTypes*, focused on how many types of time schools used to opportunities for content-focused professional development, did not have statistically significant impacts on any of the study outcomes. A final predictor, *InstrBudget,* a school's total annual budget, was omitted from analysis due to large amounts of missing data from the Matrix A sample.

The present study examined the degree to which aspects of the *Contexts for Teachers' Learning* framework (National Academies of Sciences, Engineering, And Medicine's Committee on Strengthening Science Education, 2015) appeared to impact middle grades stem teachers' instructional objectives, practices, and instructional technology use. Statistically significant school- and teacher-level predictors in all four areas of the framework suggest that the building teachers' professional capacity, coherent instructional guidance, leadership, and providing adequate time and funding are impactful areas of study for those seeking to examine stem instructional reform in the middle grades. Efforts to build teachers' *professional capacity*, including a school-level focus on developing external partnerships, as well as building teachers' feelings of both content and instructional preparedness, appear to result in positive outcomes for stem teachers' frequency of instructional technology use, use of reform-based instructional practices, and focus on reform-based instructional objectives. However, this study's results also show that efforts to build teachers' professional capacity may have differential impacts by years of experience, with more experienced teachers sometimes less likely to benefit from capacity-building efforts, specifically teachers' use of reform-based practices. A recent study of the professional development of early career teachers (Gabriel, 2010) posited that differentiated professional development could provide teachers with opportunities to focus their professional learning in ways that would help them grow more efficiently as instructors and may contribute to teacher retention in the field.

Regarding schools' efforts to facilitate *coherent instructional guidance* and teacher *leadership*, the present study's results showed that providing teachers with opportunities to learn new practices through instructional coaching and teacher participation in leadership activities, including coaching and mentoring inservice and preservice teachers, appear to positively impact both reform-based instructional objectives and practices. However, it also appears that a teachers' degree of direct participation in leadership and professional development activities rather than simply working in a school where these opportunities are available, is what results in changes to practice. These findings mirror those of a recent study of the importance of campus leadership supporting teacher participation in leadership campus-wide in Chicago Public Schools. Of the 12 schools studied in depth, those with a collaborative teacher leadership culture supported by campus administration were more likely to show academic gains for students (Allensworth & Hart, 2018). Finally, the provision of adequate *time and funding* appear to have significant impacts on middle grades STEM teachers' instructional practices and instructional technology use, although these findings were somewhat conflicting between mathematics and science teachers. In the case of reform-based practices in both mathematics and science, teachers' perceived adequacy of equipment and supplies positively predicted their use of reform-based practices. However, teachers' perceptions of technology barriers had differential impacts on mathematics and science teachers, with mathematics teachers' perceptions of a lack of technological barriers having an unexpected negative impact on both reform-based instructional practices and their frequency of instructional technology use. Science teacher perceptions of their lack of technological barriers, on the other hand, positively predicted their frequency of instructional technology use. It is likely that there is an underlying factor not explained by the model sample that is explaining the variation in the mathematics teacher responses.

Limitations of Study and Future Research

The present study, though important in its examination of how contexts for teachers' learning impact teachers' use of reform-based STEM practices, has several limitations. The study's greatest limitation is that the data 2012–13 NSSME questionnaires is about five years old. In addition, due to the similarity of some of the school- and teacher-level predictors, it is possible that one or more of the predictors functioned as a moderator of another predictor. For example, in the case of the negative impact of teachers' perceptions of a lack of technological barriers on reform-based teaching practices, it is possible that another predictor, such as years of teaching experience was moderating this relationship negatively due to the fact that over 50% of the full and sub-sample of teachers had 10 or more years of experience. Finally, the present study's focus on malleable school- and teacher-level factors did not account for fixed factors, such as socio-economic status of students and students' prior achievement level. Though this study's purpose was to serve as a preliminary examination of the relationship between malleable factors and teacher practices, it is likely that fixed factors explain a large part of the variance in teaching practice.

This study presents several opportunities for future research. First, it would be beneficial to repeat the study with NSSME samples of elementary and high school teachers to examine whether the study outcomes hold for teachers at other levels. In addition, repeating the study with expanded statistical models that include fixed factors, such as students' prior achievement level and the influence of state standards on curriculum and teaching, would determine to what degree malleable factors impact teaching practice when fixed factors are accounted for. Finally, using these, or similar, measures on a sample of teachers for whom student achievement data could be obtained could facilitate an examination of the degree to which reform-based practices result in impacts in students' STEM outcomes. Though a great deal of research has been conducted on one or more aspects of students' middle grades STEM pipeline experiences, a majority of studies rely on student-reported data. Analyses of teacher self-reports of perceptions and practices are a critical, yet under-utilized, source of data in the examination of what works in middle grades STEM. As schools across the nation wrestle with meeting high academic standards and growing student interest in STEM, teacher perspectives and classroom practices are an important area of focus for those seeking to examine STEM instructional reform in the middle grades.

CHAPTER IV

AN EXAMINATION OF SCHOOL AND TEACHER FACTORS CONTRIBUTING TO EFFECTIVE STEM PRACTICES IN TEXAS MIDDLE GRADES SCHOOLS

Across the U. S., K-12 schools are increasingly shifting their focus to student experiences in the STEM disciplines in an effort to keep pace with the growth of a knowledge-based economy focused on process-based thinking skills, problem solving, persistence, and creativity (Bailey, Kaufman, & Subotic, 2015). As the second largest state in the U. S., Texas has a high demand for a skilled STEM workforce, with an expected increase of over 100,000 workers by 2024 (Texas Workforce Investment Council, 2015). The continued growth of the STEM workforce has intensified the national and state focus on developing K-12 students' STEM competencies. This intensified focus on students' STEM competency has also resulted in increased focus on the STEM-related knowledge and practices of K-12 STEM teachers.

The result of increased scrutiny has led to a great deal of examination, and standard-setting focused around STEM teachers' instructional practices and STEM student achievement in recent years. The present study investigated teacher and school factors that explain variation in teacher self-reports of effective STEM practices in middle grades schools in the state of Texas.

The purpose of the present study is to explore teacher and school factors that explain variation in teacher self-reports of use of effective STEM practices in middle grades schools in the state of Texas. The present study's use of teacher self-report data contributes to the present literature on effective STEM instructional practices in several ways: (a) the use of a mixed-methods approach focused on teacher perceptions of their own STEM instructional practices, (b) an analysis of how school contextual factors contribute to STEM instruction, (c) an explicit focus on the middle grades, and (d) a comparison of effective and less-effective schools. Though there exists a substantial body of research on the effectiveness of STEM instructional practices on students' achievement and attitudes in STEM, few studies focus on the degree to which teachers perceive that their schools and classrooms are providing students with research-based STEM practices, as well as the degree to which school contextual factors, such as professional learning, time, and funding, influence STEM instruction. The present study's inclusion of quantitative and qualitative data from both high- and low-performing schools allows for an in-depth look into middle grades STEM classrooms across the state of Texas. Analyses of teacher self-reports of perceptions and practices are a critical, yet under-utilized, source of data in the examination of what works in middle grades STEM.

Survey questions were adapted from the National Survey of Mathematics and Science Educators (NSSME, Weis & Banilower, 2014). The NSSME was utilized due to its focus on U. S.–based K-12 STEM teachers, while the SASS provided items focused on teacher retention and job satisfaction not addressed on the NSSME. The analysis centered around teachers' perceptions of reform-based instructional objectives, instructional practices, and use of instructional technology. The following research questions guided the study:

1. What is the effect of malleable school and teacher factors on Texas middle grades STEM teachers' use of effective STEM instructional objectives, controlling for school and teacher factors?
2. What is the effect of malleable school and teacher factors on Texas middle grades STEM teachers' use of effective STEM instructional practices, controlling for school and teacher factors?
3. What is the effect of malleable school and teacher factors on Texas middle grades STEM teachers' use of instructional technology, controlling for school and teacher factors?
4. Are there differences in the use of effective STEM practices by school type (Gold Ribbon middle schools and non-Gold Ribbon middle schools)?
5. How do Texas middle grades STEM teachers' perceptions of the most effective means of teaching STEM subjects, as well their perceptions of barriers to effective STEM instruction, align with previous research?

Conceptual Framework

The conceptual framework for the present study is adapted from Hansen's (1996) STEM pipeline factors, including achievement, attitude, access, and activities, and also takes into account more recent work on impactful mechanisms for teacher and school change (Allensworth, Ponisciak, & Mazzeo, 2009; Brophy, Klein, Portsmore, & Rogers, 2008; Bryk & Schneider, 2002; DeAngelis & Presley, 2011; Maltese & Tai, 2011), including school culture (i.e., level of collaboration, classroom autonomy), professional learning (i.e., amount and types of professional development in STEM), and specific instructional practices (i.e., classroom discourse, hands-on learning experiences, real world connections). The overarching purpose of the conceptual framework is to provide a lens through which K-12 school stakeholders might examine the impact of school- and teacher-level factors on students' STEM experiences in order to increase positive STEM outcomes for students. Figure 4.1 provides an overview of the conceptual framework of the study.

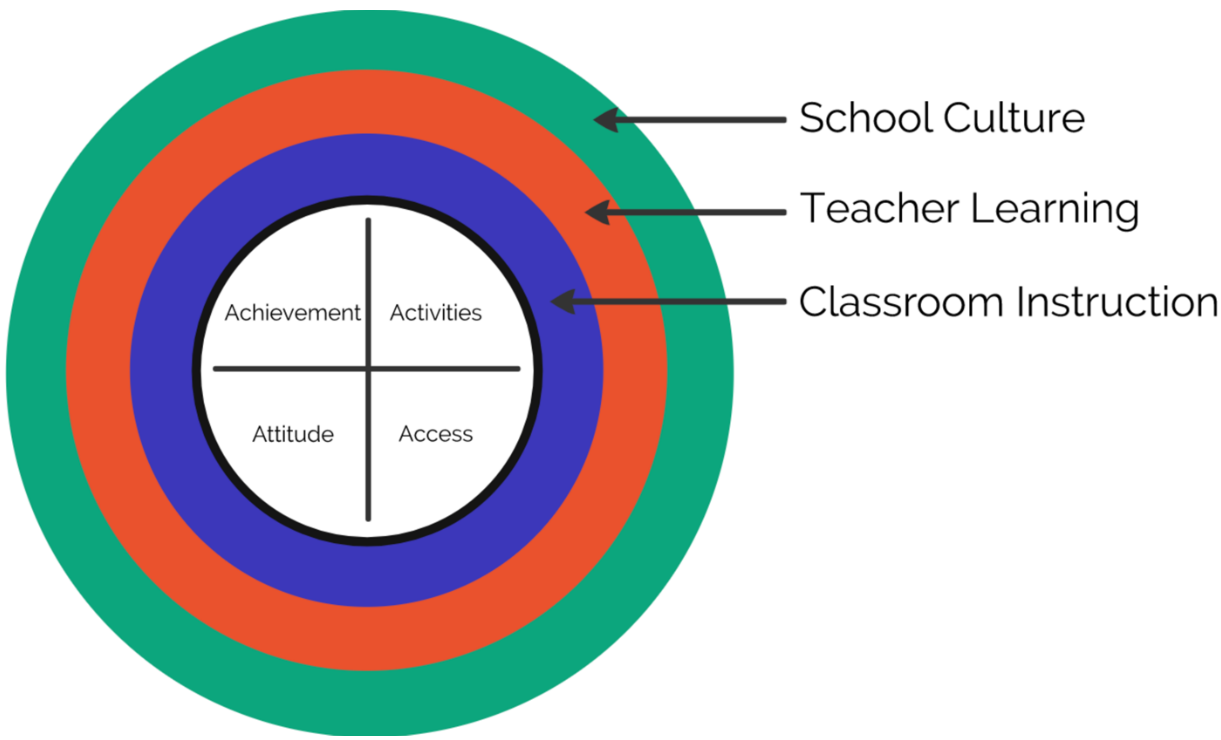


Figure 4.1. *A Statewide Examination of the Impact of School and Classroom Factors on Middle Grades Students' STEM Pipeline Experiences.*

Method

Data Sources and Sample

The Texas Middle Grades STEM Teacher Survey was used to examine middle grades STEM teachers' professional learning experiences and instructional practices, as well as school-level factors, such as school facilities and availability of technology. Open-ended questions on the survey allowed teachers to discuss what they felt were the most effective means of teaching STEM subjects, as well their perceptions of barriers to effective STEM instruction in Texas. The survey items were adapted from the National Survey of Science and Mathematics Educators Mathematics and Science Teacher Questionnaires (NSSME, Weis & Banilower, 2014) and the 2011-12 Schools and Staffing Survey (SASS) Teacher Questionnaire. The Texas Middle Grades STEM Teacher Survey was administered as an online survey in the spring of 2018. Teachers were sent an email invitation detailing the study's purpose and time required to take the survey. The email invitation provided a link to a study information sheet and online survey in the Qualtrics® online survey system. Participants choosing to complete the survey answered a question providing online consent. Teachers who did not provide online consent were directed out of the online survey. All linkages between participant names/email addresses were coded and separated from any data containing responses. In an attempt to increase survey response rates, teachers who opted to provide their name and email address were entered into a drawing for one of 15 $150 gift cards randomly drawn from participants who respond within the first 10 days of the initial invitation. Any participant not selected during the first drawing, along with any respondent who completed the survey before it closed was entered into an additional drawing of 15 $75 gift cards randomly drawn from non-selected early responders and later responders who provide their contact information.

The study's sample included science, technology, engineering, and mathematics teachers from a sample of Texas public middle schools serving 75% or higher free- and reduced-lunch students in grades six through eight. The sample list of schools was obtained from the Children At-Risk 2017 ranking of every public middle school in the state of Texas. The 47 schools identified as Gold Ribbon Schools in the Children At-Risk rankings (those schools with 75% or greater economically disadvantaged students and a high level of student performance) were propensity score matched on demographic variables with a sample of 47 non-Gold Ribbon middle schools. Each year, Children At Risk, a Texas-based non-profit organization focused on improving the quality of life for children through research, public policy analysis, education, collaboration and advocacy (Children At Risk, 2017), ranks each elementary, middle, and high school in the state of Texas based on four factors: (1) student performance on the State of Texas Assessment of Academic Readiness (STAAR) Reading and Math tests, (2) a school's overall campus performance compared to other campuses statewide with similar levels of poverty, (3) student-level improvement over time on standardized test scores in Reading, English, and Math, and (4) and the high school graduation rates, SAT/ACT participation rate and scores, and AP/IB participation rate and scores (for high schools only). Each school is assigned a letter grade from A-F based on their rankings in the total list of schools at each level (elementary, middle, and high school). Additionally, Children At Risk provides a list of "Gold Ribbon Schools," that are designated as such because they receive a letter grade of A or B and have a student population of 75% or more economically disadvantaged students (Children At Risk, 2017).

The schools in the study's initial sample consisted of Gold Ribbon middle schools that were propensity score matched based on school size and student demographic variables, such as percent free and reduced lunch and race/ethnicity, with a sample of all Texas middle schools with 75% or greater economically disadvantaged students that receive a letter grade of D or F. Schools receiving a designation of D or F received a composite score on the four ranking factors that fell below the 35th percentile of all middle schools in the state (Sanborn, Canales, Everitt, McClendon, McConnell, O'Quinn, & Treacy, 2017). The initial sample included 979 STEM teachers in grades six through eight. Fifth-grade teachers were excluded from the sample due to the fact that Children At Risk defines middle grades as grades six through eight. Participant email addresses were obtained for all mathematics, science, technology, and engineering teachers from publicly available information on school websites. Schools and teachers for whom email addresses could not be obtained were eliminated from the sample.

The sample to whom email invitations were sent included 56 campuses from 40 school districts across the state of Texas. About half of the initial sample consisted of Gold Ribbon campuses ($n = 29$) and the remaining campuses were non-Gold Ribbon ($n = 27$). The survey was launched in January of 2018. Emails from eight individual teachers and nine school districts were either blocked by a district server or bounced back as no longer active addresses. In addition to the initial email invitation, three email reminders were sent for respondents who had not completed the survey. Despite multiple email requests and the offer of gift cards for survey completion, only 115 of the 956 teachers sampled responded to the survey, for a response rate of 11.99%. The low response rate was likely due to subject areas of the sample and the time of year in which the survey was sent out. Grades six through eight mathematics teachers and eighth grade science teachers spend a great deal of time preparing for state assessments in the spring semester and were perhaps less likely to respond to a survey. The study's final sample included 115 teachers from 36 campuses, 21 of which were Gold Ribbon campuses. Just over half of the respondents (*n* = 68, 59.1%) were from Gold Ribbon campuses. Table 4.1 provides demographic information regarding the sample.

Table 4.1

Descriptive Statistics for Texas Middle Grades STEM Teacher Survey

	Texas Middle Grades STEM Teachers (*n* = 115)	
	n	%
Teaching experience		
Less than 4 years	32	28.07
5 – 9 years	28	24.56
10 – 14 years	30	26.32
15 or more years	24	21.05

Table 4.1 Continued

	n	%
Primary subject area		
Science	51	44.74
Mathematics	49	42.98
CTE/Engineering	14	12.28
Grade level		
Sixth grade only	28	24.78
Seventh grade only	26	23.01
Eighth grade only	33	29.20
Multiple grade levels	25	22.12
Type of teacher preparation program		
Traditional certification	48	42.86
Non-traditional certification	64	57.14
Campus Gold Ribbon status		
Gold Ribbon	68	59.10
Non-Gold Ribbon	47	41.90

Source. Texas Middle Grades STEM Teacher Survey

Data Analysis

The main purpose of the present study was to examine the impact of malleable factors for STEM improvement that are advocated for by both the Institute for Education Sciences, the U. S. Department of Education, and the National Science Foundation (2013) and the National Academies of Sciences, Engineering, and Medicine (2015). As such, the study's predictors focused on malleable school and teacher factors that have been identified as impactful improving teachers' STEM instructional practices. The predictors are mainly derived from the *Contexts for Teachers' Learning* framework (National Academies of Sciences, Engineering, and Medicine's Committee on Strengthening Science Education, 2015). The framework is based on a synthesis school- and teacher-level factors that have been shown to result in changes in reform-based teaching practice. Based on the work of Bryk and colleagues at the Chicago Consortium for School Reform (Bryk, Sebring, Allensworth, Suppescu, & Easton, 2010), as well as multiple studies of comprehensive school reform efforts (Cohen & Hill, 2000, 2001; Cohen, Raudenbush, & Ball, 2003; Le et al, 2006; Rowan, Corenti, Miller, & Camburn, 2009), the four main aspects of the framework are: (1) professional capacity, (2) coherent instructional guidance, (3) leadership, and (4) time and funding. The present study explored three of the four aspects of the framework, including professional capacity, coherent instructional guidance, and time and funding.

The survey contained seven sections: (1) teaching background and assignment, (2) professional development, (3) feelings of preparedness, (4) available resources for teaching, (5) perceptions of teaching and work environment, (6) demographic information, and (7) open-ended items addressing teachers' views on effective STEM instruction and barriers to high-quality STEM instruction. The survey was piloted with 23 middle grades STEM teachers in December of 2017. Feedback from teachers and descriptive statistics of items were used to revise the survey. The survey was reduced from 66 to 44 total questions. In addition to the teacher pilot, the survey was shared with a team of STEM education and education research content experts to establish the content validity of survey items.

The initial plan for data analysis was hierarchical linear modeling to account for the nesting of teachers within schools. However, due to the small sample size, multiple linear regression was used to examine the relationship of study predictors on target outcomes. The survey's two open-ended questions allowed teachers to discuss what they felt were the most effective means of teaching STEM subjects, as well their perceptions of barriers to effective STEM instruction in Texas. Constant comparative analysis of teachers' responses (Glaser & Strauss, 1967) was utilized to determine themes from the open-ended responses. Due to a large number of responses, each of the responses was quantified by theme in order to provide a frequency count of the number of responses mirroring each of the emergent themes (Tashakkori & Teddlie, 1998). In addition, quotes or phrases illustrative of each theme are provided to expand the description of each theme. Finally, as a means of triangulating qualitative responses with quantitative data, the emergent themes for barriers to high quality STEM instruction were examined for the degree to which they related to the four *a priori* elements of the *Contexts for Teachers' Learning* framework (National Academies of Sciences, Engineering, and Medicine's Committee on Strengthening Science Education, 2015), while the emergent themes for instructional methods were examined for the degree to which they related to the NSSME Reform-Oriented Instructional Objectives and Teaching Practices (Weis & Banilower, 2014).

Dependent Measures

The three dependent variables focused on the following: (1) teachers' instructional objectives, (2) instructional practices, and (3) frequency of instructional technology use. The three instructional areas are seen as influential in generating and increasing students' interest and achievement in STEM (Brotman & Moore, 2008; Brown, Concannon, Marx, Donaldson, & Black, 2016; Christensen, Knezek, & Tyler-Wood, 2015; Dare & Roehrig, 2016; Hansen & Gonzalez, 2014; Kim, 2016; Nugent, Barker, Welch, Grandgenett, Wu, & Nelson, 2015; Riegle-Crumb, Moore, & Ramos-Wada, 2010). The dependent measures were adapted from the NSSME Mathematics Teacher Questionnaire (MTQ) and Science Teacher Questionnaire (STQ) that were common to both mathematics and science. Career and Technology Education (CTE) and other STEM teachers did not receive these items, but responded to open-ended items focusing on best STEM instructional practices and barriers to STEM instruction.

Principal-components factor analyses with Varimax rotation using Stata 15 statistical analysis software was conducted on items for each dependent measure. Eigenvalues greater than 1.00 were used to determine the number of factors, with each item's highest factor loading determining its scale. A regression-based factor score was predicted from the items on each scale. Cronbach's alpha was calculated to determine the internal consistency of the items composing each scale. The results of the factor analysis, factor loadings, and scale reliabilities for mathematics and science are shown in Tables 16 – 21. Factors with Cronbach's alpha reliabilities of less than 0.65 were dropped from analysis due to low reliabilities (Loewenthal, 2001). The third dependent measure, teachers' use of instructional technology, was not retained due to less than acceptable Cronbach alpha reliabilities for all retained factors.

Mathematics and Science Instructional Objectives. Mathematics and science survey respondents were asked to indicate how much emphasis they placed on seven instructional objectives over the course of the school year. All items were scored on a 4-point Likert-type measure with 1 = No emphasis, 2 = Minimal emphasis, 3 = Moderate emphasis, and 4 = A great deal of emphasis. The analysis yielded two factors with Eigenvalues above 1.00, accounting for 63.38% of the variance. Item factor loadings ranged from 0.248 to 0.831. Factor 1 was labeled Reform-Based Objectives (RBO), focusing mainly on conceptual understanding and increasing student understanding, while Factor 2, labeled Procedurally-focused Objectives (PFO), contained two items focused on procedural understanding and test preparation. In addition, the internal consistency reliability (Cronbach alpha) of each scale was calculated. The internal consistency reliability coefficients of the scales were 0.50 to 0.79, with an acceptable alpha level for the RBO factor only. The PFO factor was not retained for analysis due to its low reliability. Table 4.2 exhibits the items and their corresponding factor loadings.

Table 4.2

Mathematics and Science Instructional Objectives Items and Factor Loadings

	Factor Loadings	
	RBO	PFO
Understanding concepts	0.771	
Disciplinary practices	0.573	
Real-life applications	0.807	
Increasing student interest in subject	0.809	
Preparing for further study	0.831	
Procedures/Memorization/Algorithms		0.770
Learning test-taking skills/strategies		0.746
Eigenvalue	3.24	1.20
Cronbach's alpha reliability	0.79	0.58
Total variance explained by factors	63.38%	

Mathematics and Science Instructional Practices. Survey respondents were asked to indicate the frequency with which they focused on 10 instructional practices, including items ranging from *engaging the whole class in discussions* to *having students work in small groups*. All items were scored on a 4-point Likert-type measure with 1 = Never, 2 = Rarely (a few times a year), 3 = Sometimes (once a month), and 4 = Often (daily or weekly). The analysis yielded three factors with Eigenvalues above 1.00, accounting for 60.63% of the variance. Item factor loadings ranged from 0.502 to 0.815.

Factor 1 was labeled Reform-Based Practices (RBP), focusing mainly on practices related to small group work and student reflection. Factor 2, labeled Whole Class Instruction (WCI) focused on whole group instruction and discussion. Factor 3 (Literacy and Assessment Practices [LAP]) focused on reading from a textbook and test preparation. The internal consistency reliability coefficients of the scales ranged from 0.52 to 0.75, with an acceptable alpha level for the RBP and WCI factors only. The LAP was not retained for analysis due its low reliability. Table 4.3 exhibits the items and their corresponding factor loadings.

Table 4.3

Mathematics and Science Instructional Practices Items and Factor Loadings

	Factor Loadings		
	RBP	WCI	LAP
Have students work in small groups	0.527		
Guest lectures focusing on content in the STEM workplace	0.641		
Focus on literacy skills	0.723		
Give tests/quizzes that include constructed-response/open-ended items	0.701		
Have students write reflections	0.815		
Explain ideas to the whole class		0.857	
Engage the whole class in discussions		0.888	

Table 4.3 Continued

	Factor Loadings		
	RBP	WCI	LAP
Read from textbook or other material			0.502
Have students practice for standardized tests			0.751
Give tests/quizzes that are predominately short answer			0.664
Eigenvalue	3.31	1.57	1.18
Cronbach's alpha reliability	0.75	0.74	0.52
Total variance explained by factors	60.63%		

Mathematics and Science Instructional Technology Use. Survey respondents were asked to indicate the frequency with which they used on nine types of instructional technology, including laptops and various types of calculators. All items were scored on a 4-point Likert-type measure with 1 = Never, 2 = Rarely (a few times a year), 3 = Sometimes (once a month), and 4 = Often (daily or weekly). The analysis yielded four factors with Eigenvalues above 1.00, accounting for 68.75% of the variance. Item factor loadings ranged from 0.476 to 0.930. Factor 1, student-centered technology (SCT) included handheld devices, such as clickers and smart phones. Factor 2, routine technology (RT), included personal computers and internet. Factors 3 and 4 included calculators (CALC) and tablets (TAB). Due to low reliabilities for all factors, the instructional technology use dependent measure was not retained for analysis. Table 4.4 exhibits the items and their corresponding factor loadings.

Table 4.4

Instructional Technology Use Items and Factor Loadings

	Factor Loadings			
	SCT	RT	CALC	TAB
Four-function calculators	0.792			
Probes for collecting data	0.697			
Classroom response systems or clickers	0.476			
Smart phones	0.578			
Personal computers, including laptops		0.897		
Internet		0.669		
Scientific calculators			0.930	
Graphing calculators			0.702	
Tablets				0.859
Eigenvalue	2.37	1.44	1.23	1.14
Cronbach's alpha reliability	0.57	0.55	0.64	
Total variance explained by factors	68.75%			

Independent Measures

The present study's predictors explored three of the four aspects of the *Contexts for Teachers' Learning* framework (National Academies of Sciences, Engineering, and Medicine's Committee on Strengthening Science Education, 2015), including professional capacity, coherent instructional guidance, and time and funding. *Professional capacity focuses* on efforts by teachers and schools to build the instructional and collaborative capacity of staff through a variety of means, including collaboration, staff qualifications, and partnerships. The two predictors included teachers' perceptions of their level of content preparedness (*ContentPreparedness*), and teachers' perceptions of their level of instructional preparedness (*InstrPreparedness*). The second aspect of the framework, *coherent instructional guidance*, focuses on the degree to which teachers have opportunities to learn new practices and consider ways in which new practices might be adapted for successful implementation in their classroom and school contexts. Three predictors addressed coherent instructional guidance. The predictors included the extent to which a teacher had participated in content-specific professional development in the last three years (*ContentPDType*), the degree to which the professional development aligned with research-based aspects of high-quality STEM professional development (*PDEmphasis*), and a final predictor focused on the extent to which professional development content focused on various aspects of student-centered instruction (*StudentFocusedPD*). The final aspect of the framework examined two predictors related to the allocation of *time and funding*. The two predictors focused teachers' perceptions on the availability of equipment and supplies (*ResourceAdequacy*) and availability of different types of instructional technology (*TechAvailability*).

In addition to the malleable factors, school and teacher control variables included whether a school was a Gold Ribbon campus, teaching experience, subject area, grade level, type of teacher preparation program, whether or not a teacher had a STEM degree, and students' prior achievement. Three control variables, teacher sex, whether a teacher identified as Hispanic/Latinx, and race/ethnicity were eliminated from analysis due to large amounts of missing data. Categorical survey items were analyzed as categorical data, while several numeric items, such as years of teaching experience, were banded into categories to examine commonalities within sub-groups of teachers. The remaining predictors were composite variables of sub-items addressing similar constructs. For ease of interpretation, all composite variables were scaled by summing values across all items in a scale and dividing by the total possible value for all items, assigning each composite variable a value between zero and 1. Table 4.5 provides information regarding the data type and descriptive statistics of each of the predictors used. The mean and standard deviation is given for all scaled items and the number and percent of responses in each category is provided for categorical and dichotomous items.

Table 4.5

School and Teacher Predictors – Data Types and Sample Descriptive Statistics

	Data Type	Descriptive Statistics	
Professional Capacity		*M/n*	*SD/%*
ContentPreparedness	Scale	0.90	0.11
InstrPreparedness	Scale	0.81	0.14
Coherent Instructional Guidance		*M/n*	*SD/%*
ContentPDType	Dichotomous	0.64	0.28
PDEmphasis	Dichotomous	0.73	0.17
StudentFocusedPD	Scale	0.69	0.16
Time and Funding		M/n	SD/%
ResourceAdequacy	Scale	0.83	0.11
TechAvailability	Dichotomous	0.68	0.15
Control Variables			
GoldRibbonCampus	Dichotomous		
Yes		29	51.80
No		27	48.20
Teaching experience	Categorical		
Less than 4 years		32	28.07

Table 4.5 Continued

	Data Type	Descriptive Statistics	
5-9 years		28	24.56
10-14 years		30	26.32
15 or more years		24	21.05
Grade level	Categorical		
Sixth grade only		28	24.78
Seventh grade		26	23.01
Eighth grade		33	29.20
More than 1 grade level		25	22.12
STEM Subject			
Mathematics		49	42.98
Science		51	44.74
CTE/Engineering		14	12.28
Teacher preparation program	Categorical		
Traditional certification		48	42.86
Non-traditional certification		64	57.14
STEM degree	Dichotomous		
Yes		50	43.48%
No		65	56.52%

Table 4.5 Continued

	Data Type	Descriptive Statistics	
Students' prior achievement	Categorical		
Mostly low achievers		27	26.47
Mostly average achievers		18	17.65
Mostly high achievers		9	8.82
A mixture of all levels		48	47.06

Source. Texas Middle Grades STEM Teacher Survey

Results

STEM Instructional Objectives

A multiple linear regression was calculated to predict teachers' use of effective STEM instructional objectives based on school and teacher malleable predictors (technology availability, content preparedness, instructional preparedness, number of types of professional development attended, content of professional development, emphasis of professional development, and perception of adequacy of resources) and control variables (Gold Ribbon campus teaching experience, subject area, grade level, type of preparation program, STEM degree, and prior achievement level of students). Listwise deletion was used for with missing data on the included measures (Enders, 2010). A significant regression equation was found (F(21, 58= 3.46, $p < .001$) with an R^2 of 0.556. Only one malleable predictor significantly predicted teachers' use of reform-based science instructional objectives. As teachers' perceptions of their level of instructional preparedness (*InstrPreparedness*) increased, their use of reform-based objectives increased by 3.79 points on average. In addition, teachers with 10 – 14 years of teaching experience had a reform-based instructional objectives score that was 0.60 points higher, on average, than teachers with fewer than five years of experience. Finally, mathematics teachers had a reform-based instructional objectives score that was 0.58 points lower, on average, than science teachers. Table 4.6 provides the multiple regression results for teachers' perceptions of reform-based objectives. CTE and Engineering teachers did not answer content preparedness questions and were therefore, were deleted from the analyses.

Table 4.6

Results of Multiple Regression Analysis –Reform-based Instructional Objectives

Predictor	*b*	SE	*t*	*p*	*95% CI*	
Malleable predictors					Lower Bound	Upper Bound
ContentPreparedness	1.191	1.029	1.160	0.252	-0.869	3.251
InstrPreparedness	3.787	0.947	4.000	0.001**	1.892	5.681
ContentPDType	0.237	0.379	0.630	0.534	-0.522	0.997
PDEmphasis	0.250	0.882	0.280	0.778	-1.516	2.015
StudentFocusedPD	0.516	0.829	0.620	0.536	-1.143	2.175

Table 4.6 Continued

Predictor	*b*	SE	*t*	*p*	*95% CI*	Predictor
ResourceAdequacy	0.935	1.122	0.830	0.408	-1.311	3.181
TechAvailability	0.090	0.747	0.120	0.905	-1.405	1.584
Control Variables						
GoldRibbonCampus	0.394	0.234	1.680	0.098	-0.074	0.862
Teaching experience						
5-9 years	0.495	0.298	1.660	0.102	-0.102	1.091
10-14 years	0.601	0.286	2.100	0.040*	0.028	1.175
15 or more years	0.177	0.303	0.580	0.562	-0.430	0.784
Grade level						
Seventh grade	-0.465	0.305	-1.530	0.132	-1.075	0.145
Eighth grade	0.065	0.263	0.250	0.807	-0.462	0.591
More than 1 grade level	-0.283	0.305	-0.930	0.357	-0.893	0.327
Primary Subject area†	-0.581	0.246	-2.360	0.022*	-1.074	-0.087

Table 4.6 Continued

Predictor	*b*	SE	*t*	*p*	*95% CI*	Predictor
Teacher preparation program	0.041	0.204	0.200	0.840	-0.366	0.449
STEM degree	-0.029	0.254	-0.110	0.910	-0.538	0.480
Students' prior achievement						
Mostly average achievers	0.279	0.289	0.960	0.339	-0.300	0.858
Mostly high achievers	0.458	0.393	1.170	0.248	-0.328	1.244
A mixture of all levels	0.064	0.251	0.260	0.799	-0.439	0.568

Note. *, $p < .05$** $p < 0.01$.

STEM Instructional Practices

Two multiple linear regressions were calculated to predict teachers' use of instructional practices. The first factor, reform-based practices (RBP), focused mainly on practices related to small group work and student reflection, while the second factor, teachers' use of whole class instruction and discussion (WCI), focused on the extent to which teachers used whole group instruction and engaged their students in whole class discussions. Teachers' scores for both factors were based on school and teacher malleable predictors (technology availability, content preparedness, instructional preparedness, number of types of professional development attended, content of professional development, emphasis of professional development, and perception of adequacy of resources) and control variables (Gold Ribbon campus teaching experience, subject area, grade level, type of preparation program, STEM degree, and prior achievement level of students). A significant regression equation was found for the first instructional practice factor, teachers' use of reform-based practices, (F(21, 56= 2.08, $p < .05$) with an R^2 of 0.437. The regression equation for the second factor, whole class instruction, was not significant (F(21, 56= 0.95, $p = .533$). Two malleable predictors significantly predicted teachers' use of reform-based instructional practices. As teachers' perceptions of their level of instructional preparedness (*InstrPreparedness*) increased, their use of reform-based practices increased by 2.15 points on average. In addition, as teachers' perception of the availability of instructional technology increased, their use of reform-based instructional practices increased by 1.92 points on average. None of the control variables significantly predicted teachers' use of reform-based instructional practices. Table 4.7 provides the multiple regression results for teachers' perceptions of their use of reform-based instructional practices.

Table 4.7

Results of Multiple Regression Analysis –Reform-based Instructional Practices

Predictor	*b*	SE	*t*	*p*	*95% CI*	
Malleable predictors					Lower Bound	Upper Bound
TechAvailability	1.916	0.865	2.22	0.031*	0.184	3.649
ContentPreparedness	-0.473	1.195	-0.40	0.694	-2.868	1.922
InstrPreparedness	2.150	1.038	2.07	0.043*	0.070	4.230
ContentPDType	0.367	0.418	0.88	0.383	-0.470	1.204
PDEmphasis	-0.006	0.958	-0.01	0.995	-1.926	1.913
StudentFocusedPD	0.666	0.912	0.73	0.468	-1.160	2.493
ResourceAdequacy	-0.474	1.266	-0.37	0.709	-3.009	2.06
Control Variables						
GoldRibbonCampus	0.113	0.257	0.440	0.662	-0.401	0.626
Teaching experience						
5-9 years	0.417	0.333	1.250	0.215	-0.250	1.085

Table 4.7 Continued

Predictor	*b*	SE	*t*	*p*	*95% CI*	
10-14 years	0.168	0.321	0.520	0.604	-0.476	0.812
15 or more years	-0.395	0.337	-1.17	0.247	-1.070	0.281
Grade level						
Seventh grade	-0.059	0.340	-0.170	0.863	-0.740	-0.740
Eighth grade	-0.295	0.290	-1.020	0.314	-0.876	-0.876
More than 1 grade level	0.270	0.337	0.800	0.427	-0.406	0.622
Primary Subject area†	-0.396	0 .283	-1.40	0.167	-0.962	0.170
Teacher preparation program	0.023	0.230	0.100	0.922	-0.439	0.484
STEM degree	-0.132	0.279	-0.470	0.639	-0.691	0.428

Table 4.7 Continued

Predictor	*b*	SE	*t*	*p*	*95% CI*	
Students' prior achievement						
Mostly average achievers	-0.073	0.313	-0.230	0.817	-0.699	0.554
Mostly high achievers	-0.276	0.428	-0.650	0.521	-1.134	0.581
A mixture of all levels	-0.071	0.278	-0.250	0.800	0.554	0.487

Note. * $p < .05$**, $p < 0.01$, † = CTE/Engineering teachers did not answer content preparedness questions and were deleted from the analyses.

Instructional Technology Use

Due to low reliabilities for all instructional technology use factors, the instructional technology use dependent measure was not retained for analysis.

Use of Effective STEM Practices by School Type

A final quantitative research question focused on whether there were instructional differences between Gold Ribbon and non-Gold Ribbon campuses. Due to the nature of how the Gold Ribbon status is assigned by Children At Risk, with Gold Ribbon schools performing in the top two quintiles of all middle schools in the state of

Texas while serving high percentage of low socio-economic students, researchers were seeking to determine whether there were instructional differences between the Gold Ribbon schools surveyed and the propensity-score matched sample of non-Gold Ribbon schools, who were performing near the bottom of all middle schools in the state of Texas with similar student demographics. A one-way multivariate analysis of variance (MANOVA) by Gold Ribbon status (Gold Ribbon campus or non-Gold Ribbon campus) was performed on the dependent variable measures (teachers' reform-based instructional objectives and reform-based instructional practices) to determine if there were any differences by group. The results of the overall MANOVA did not reveal a significant difference between the two groups (*Wilks' lambda* = .985, $F(1, 82) = 0.62$, $p = .539$) on teachers' perceptions of their use of reform-based instructional objectives or instructional practices.

Teacher Perceptions of Barriers to STEM Instruction and High-Quality Instruction

The final research question explored teachers' perceptions of barriers to high-quality STEM instruction, along with perceptions of effective strategies for STEM instruction. Teachers' responses to the two open-ended survey questions were qualitatively analyzed with a focus on emergent themes. Responses that included multiple themes were coded accordingly, resulting in a total *n* that was greater than the total number of survey respondents. A total of 101 teachers responded to the survey item regarding their perceptions of barriers to high quality STEM instruction, with 106 teachers describing their perceptions of methods of effective STEM instruction.

Barriers to High-Quality STEM Instruction. Seventeen themes emerged from teacher responses regarding barriers to high-quality STEM instruction in Texas middle schools. Three themes were present in over 10% of the 143 total coded responses, including curriculum and training for teachers (15.38%), time for instruction and/or professional development (11.89%), and funding (10.49%). Regarding curriculum and training for teachers, one respondent mentioned that, "The greatest barrier would be lack of training for better STEM instruction," while another noted a need for, "research-based programs that are effective." Time barriers noted by respondents included both time for teacher collaboration and instructional time with students. One teacher expressed a need for, "planning time to create quality lessons," while another teachers' response encapsulated frustrations with a lack of time and a lack teacher training, noting, "The availability of time. We often run out of time and then we STAAR test. Also, I have never conducted half of the STEM labs myself. I would like a professional development where they teach us how to conduct the lab ourself [sic]." Other themes emerging from teachers' responses included a lack of resources, including materials and manipulatives; external expectations that were either too high, in the case of core STEM courses, such as mathematics or science, or too low, as in the case of STEM elective courses, including career and technology education; a lack of technology, mainly for students; and low student engagement and motivation. Table 4.8 exhibits each theme, along with its frequency and percentage of total responses.

Table 4.8

Emergent Themes from Teachers' Perceptions of Barriers to High-quality STEM Instruction

	Coded Teacher Responses (n = 143)	
	n	%
Curriculum/training for teachers	22	15.38%
Time	17	11.89%
Funding	15	10.49%
Resources	12	8.39%
External expectations	11	7.69%
Technology	10	6.99%
Student engagement/motivation	9	6.29%
Testing	9	6.29&
Student academic readiness	7	4.90%
Lack of administrative support	7	4.90%
Lack of space/Too many students	6	4.20%
Equipment	5	3.50%
Lack of time for collaboration	4	2.80%
Student behavior	3	2.10%
Family/parental support	3	2.10%

Table 4.8 Continued

	n	%
Teacher motivation	2	1.40%
No Barriers	1	0.70%
Total	143	100.0%

Source. Texas Middle Grades STEM Teacher Survey.
Note. Responses containing evidence of more than one theme were dual-coded resulting in an *n* that is greater than the total number survey respondents.

In addition to frequency counts of emergent themes, participants' responses were examined for the degree to which they potentially related to the four *a priori* elements of the *Contexts for Teachers' Learning* framework (National Academies of Sciences, Engineering, and Medicine's Committee on Strengthening Science Education, 2015). The National Academies document and participant responses were carefully reviewed to determine if responses in a category mirrored one of the four framework constructs. Teachers' responses and the four constructs of the framework were clearly connected in most cases. Three emergent themes, *lack of time for collaboration*, *external expectations*, and *teacher motivation* were classified as evidence of perceived weaknesses in professional capacity of teachers, with a total of 11.89% of coded responses falling into this category. Two themes, *curriculum and training for teachers* and *testing* were classified as teachers' perceptions of a lack of coherent instructional guidance due to the fact that curriculum, training, and assessment are explicitly mentioned in the framework as important aspects of coherent instructional guidance. However, though the framework focused on assessment as a necessary component of coherent instructional guidance, many respondents mentioned testing as a driving force in curricular and time constraints, as they perceived that much of their instruction was geared towards state testing. Just over 20% of coded responses fell into this category (21.68%). One theme, *lack of administrative support*, was classified as evidence of a barrier in leadership, with just under 5% of responses (4.90%) in this category. By far the largest category of responses regarding instructional barriers were classified into the time and funding category, with six emergent themes classified in this larger category. The six themes, comprising 45.45% of coded responses, included instructional time, funding, resources, technology for students and teachers, a lack of space, and overcrowding of classes due to too many students.

Four final themes, including *a lack of student engagement or motivation, student academic readiness, student behavior,* and *family or parental support*, were not addressed in the *Contexts for Teachers' Learning* framework. Just over 15% (15.38%) of survey respondents mentioned one or more aspects of student or parent relationships or attitude as a barrier to effective STEM instruction. Though not mentioned in the *Contexts for Teachers' Learning* framework, relationships among students, parents, and teachers have been found to be important for creating positive learning contexts in other recent research focused on factors that influence students' matriculation into undergraduate STEM majors. In a narrative inquiry study with graduate and undergraduate STEM students, researchers found relationships between students, teachers, and parents to be impactful in students' choice of STEM majors (Craig, Verma, Stokes, Evans & Abrol, 2018). Similarly, in their foundational work examining how school contextual factors interact for school improvement in Chicago Public Schools, Bryk and Schneider (2002) found similarly that relationships breed trust between students, teachers, and parents, without which, school improvement is unlikely. Figure 4.2 provides an overview of the overall percentage of responses in each of the four categories of the *Contexts for Teachers' Learning* framework, along with the additional category addressing student, parent, teacher relationships or attitude.

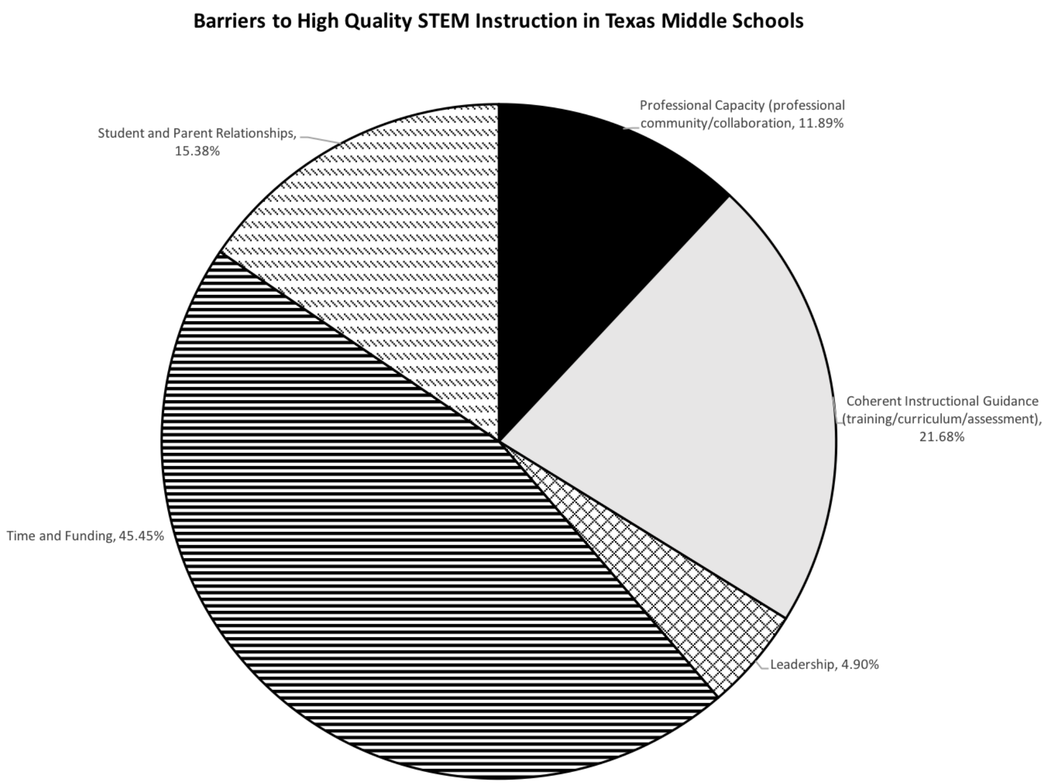


Figure 4.2. *Barriers to High Quality STEM Instruction Categorized into Contexts for Teachers' Learning Framework.*

Analysis of qualitative and quantitative data revealed several connections between teachers' open-ended responses and responses on the survey's quantitative items. For example, as teachers' perceptions of their level of instructional preparedness (*InstrPreparedness*) increased, their use of reform-based objectives increased by 3.79 points on average and their use of reform-based practices increased by 2.15 points on average. In addition, as teachers' perception of the availability of instructional technology increased, their use of reform-based instructional practices increased by 1.92 points on average. It appears that some of the same things that teachers perceive to be barriers to high quality instruction, such as access to technology and feeling prepared instructionally to do things such as planning differentiate instruction and encourage interest in STEM content for all students, including encourage females and minority students, are the same things that positively predict reform-based objectives and practices.

Methods of High-Quality STEM Instruction. Thirteen themes emerged from teacher responses regarding which methods of STEM instruction are most effective. Three themes were present in over 10% of the 200 total coded responses, including hands-on activities (20.50%), student-centered processes, including written reflections and student-led discussions (12.50%), and teacher-directed instruction, including whole group instruction and teacher-led discussion (10.50%). Regarding hands-on activities, one teacher noted that, "Hands-on activities and/or games are involved in the most effective lessons I've taught," while a science teacher expressed that, "The most effective way to teach science is through hands on activities." Student-centered processes discussed by respondents included things such as small group activities, student discussion, and writing. One teacher responded that, "Students learn best by doing and talking," while another teacher's response connected hands-on activities and student-centered processing, noting that students should have "… the opportunity to physically do something and talk about it with peers." Other themes emerging from teachers' responses included incorporating real world objects or connecting content with real world issues, specific instructional methodologies, such as inquiry learning or project-based learning, and individual drill and practice for students. Table 4.9 shows each theme, along with its frequency and percentage of total responses.

Table 4.9

Emergent Themes from Teachers' Perceptions of High-quality STEM Instructional Methods

	Coded Teacher Responses (n = 200)	
	n	%
Hands-on activities	41	20.50%
Student-centered processes	25	12.50%
Teacher-directed instruction	21	10.50%
Real-world applications	19	9.50%
Inquiry/project-based learning	18	9.00%

Table 4.9 Continued

	n	%
Engaging lessons/activities	17	8.50%
Individual student practice	14	7.00%
Appropriate pacing/time allocation	11	5.50%
Barrier in lieu of method	10	5.00%
Lab experiences/research	9	4.50%
Technology-integration	9	4.50%
Vocabulary instruction	5	2.50%
Cross-curricular integration	1	0.50%
Total	200	100.0%

Source. Texas Middle Grades STEM Teacher Survey.
Note. Responses containing evidence of more than one theme were dual-coded resulting in an *n* that is greater than the total number survey respondents.

In addition to frequency counts of emergent themes, participants' responses were examined for the degree to which they reflected reform-based practices similar to practices classified as reform-oriented in the NSSME from which the survey questions were adapted. The NSSME Public Release Datasets User Manual (Weis & Banilower, 2014) and participant responses were carefully reviewed to determine if responses in a category mirrored reform-oriented instructional objectives and practices. Teachers' responses were mirrored in the NSSME reform-oriented objectives and practices in the majority of cases, with 64.50% of responses classified as reform-oriented. Three emergent themes, *student-centered processes*, *engaging lessons/activities,* and *real-*

world applications aligned with four NSSME reform-oriented instructional objectives, including having students work in small groups, having students write their reflections in class or for homework, increasing students' interest in mathematics/science, and learning about real-life applications of mathematics/science, with a total of 30.50% of coded responses aligned to reform-based instructional objectives. Three additional themes, *hands-on activities, lab experiences/research, and inquiry- or project-based learning methods* aligned with two reform-based instructional practices, including doing hands-on/laboratory practices and engaging the class in project-based learning activities. A total of 34% of teachers' coded responses aligned to reform-based instructional practices. One theme, *technology-integration,* was not specifically identified as a reform-based practice by the NSSME, however, teachers' responses mentioning the use of instructional technology are highlighted due to the fact that access to technology for students has been found to be a limiting factor for underrepresented minority students (Aschbacher, Li, & Roth, 2010; Hill, Corbett, and St. Rose, 2010; National Research Council, 2012) and has also been found to contribute to increased student interest and achievement in STEM, either alone (Brown et al., 2016), or in combination with STEM pedagogical approaches, such as project-based learning (Hansen & Gonzalez, 2014; Kim, 2016). Just under 5% of the coded responses (4.5%) addressed the importance of technology use as a means of effective STEM instruction.

While the majority of coded responses focused on reform-oriented instructional objectives, practices, or use of instructional technology, 31% of the responses focused on instructional strategies not necessarily considered to be reform-oriented, such as *teacher* *directed instruction* or *individual student practice.* Two themes, including *vocabulary instruction* and *appropriate pacing/time allocation*, though perhaps not explicitly reform-oriented, have been shown as effective in improving STEM outcomes for students, particularly for students in middle grades mathematics and students who are English-language learners. In a 2005 study measuring the impact of extended block scheduling of sixth grade students' mathematics achievement (Biesinger, Crippen, & Muis, 2008), students at campuses with more time blocked for mathematics instruction showed statistically significant increases in mathematics achievement. In addition, two studies of the incorporation of vocabulary instruction into fifth grade science courses showed statistically significant increases in ELL students' science achievement (Lara-Alecio, Tong, Irby, & Guerrero, 2012; Llosa, Lee, Jiang, Haas, O'Connor, Van Booven, & Kieffer, 2016). Finally, 5% of teachers' responses did not focus on a pedagogical strategy, but instead focused on barriers to effective instruction or a need to teach students' behavioral strategies or compliance to teacher instructions. Figure 4.3 provides an overview of the overall percentage of responses addressing reform-oriented objectives and practices, as well as non-reform-oriented objectives and practices.

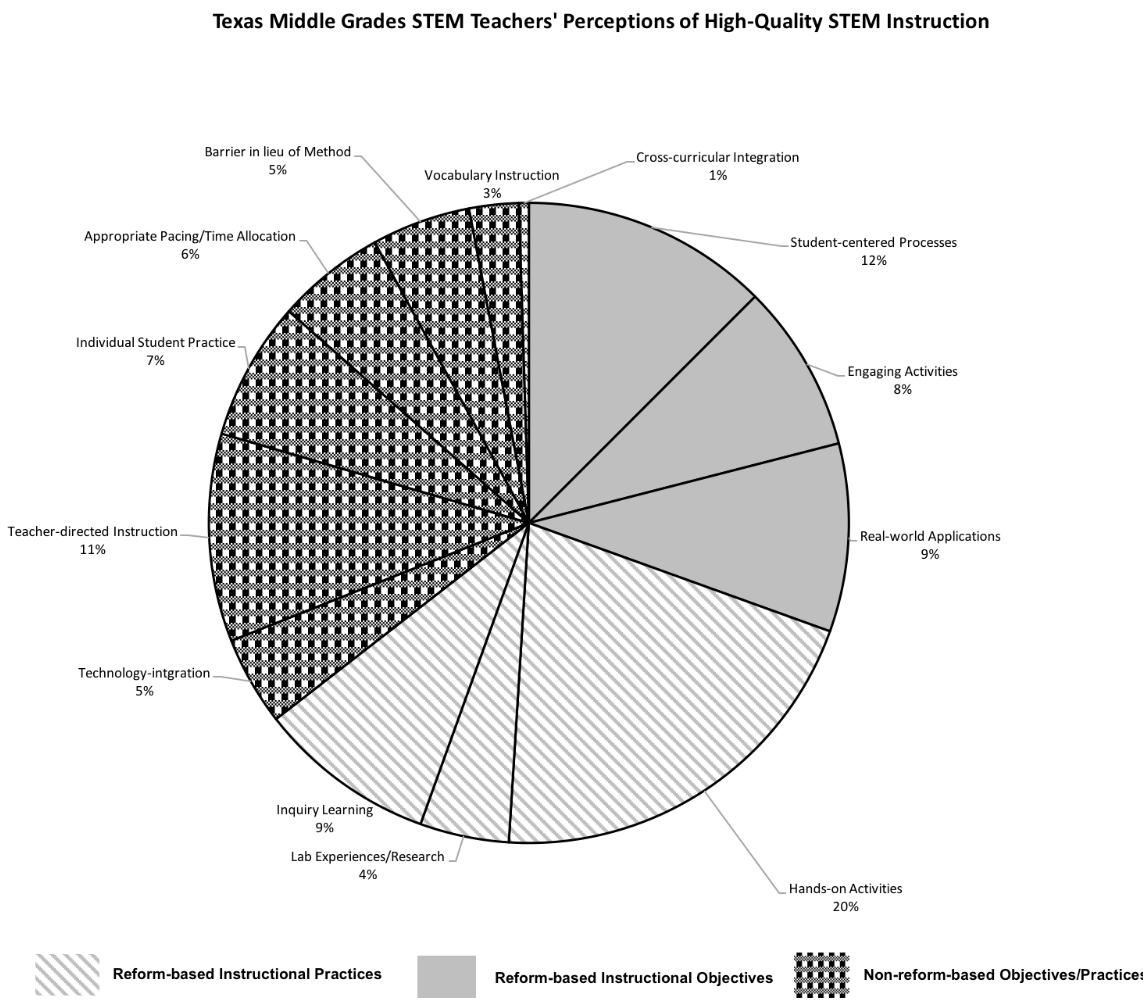


Figure 4.3. *High Quality STEM Instruction Categorized into Reform-oriented and Non-reform-oriented Practices.*

Summary and Discussion

The purpose of the present study was to explore teacher and school factors that explain variation in teacher self-reports of use of effective STEM practices in middle grades schools in the state of Texas. Survey analysis focused on teachers' perceptions of reform-based instructional objectives, instructional practices, and use of instructional technology. The main purpose of analysis was to examine the impact of malleable factors for STEM improvement drawn from the *Contexts for Teachers' Learning* framework (National Academies of Sciences, Engineering, and Medicine's Committee on Strengthening Science Education, 2015). Analysis of both qualitative and quantitative data explored three of the four aspects of the framework, including professional capacity, coherent instructional guidance, and time and funding. In addition, qualitative responses regarding teachers' perceptions of high-quality STEM instruction were analyzed for presence of absence of reform-based instructional objectives and practices classified in the NSSME from which the survey questions were adapted, while teachers' perceptions of high-quality STEM practices. The study's major findings are summarized below according to *Contexts for Teachers' Learning* framework and the NSSME reform-based practices.

Professional Capacity

Malleable aspects of building teachers' professional capacity focused mainly on efforts by teachers and schools to build the instructional and collaborative capacity of staff through a variety of means, including collaboration, staff qualifications, and partnerships. Only teachers' perceptions of their level of instructional preparedness (*InstrPreparedness*) significantly predicted teachers' use of reform-based science instructional objectives or practices. Teachers' instructional preparedness score was a composite scale measure created by summing values across all items in a scale and dividing by the total possible value for all items, assigning each composite variable a value between zero and 1. The variable included teachers' perceptions of how prepared teachers felt to use 10 types of instruction, including things such as teaching content to students with disabilities or encouraging students' interest in STEM content. As teachers' perceptions of their level of instructional preparedness (*InstrPreparedness*) increased, their use of reform-based objectives increased by 3.79 points on average. As teachers' perceptions of their level of instructional preparedness (*InstrPreparedness*) increased, their use of reform-based practices increased by 2.15 points on average. In addition to the impact of professional capacity on teachers' reform-based objectives and instruction, three emergent qualitative themes, *lack of time for collaboration*, *external expectations*, and *teacher motivation* were mentioned by teachers as barriers to high-quality STEM instruction, with a total of 11.89% of coded responses falling into this category.

Coherent Instructional Guidance

Coherent instructional guidance focuses opportunities for teachers to learn new practices and consider ways in which new practices might be adapted for successful implementation in their classroom and school contexts. Three predictors in the multiple regression equations were related to coherent instructional guidance, including the extent to which a teacher had participated in content-specific professional development in the last three years (ContentPDType), the degree to which the professional development aligned with research-based aspects of high-quality STEM professional development (PDEmphasis), and the extent to which professional development content focused on various aspects of student-centered instruction (StudentFocusedPD). None of the predictors significantly predicted any of the study's dependent measures. Two qualitative themes, *curriculum and training for teachers* and *testing* were present in over 20% of coded responses of barriers to high-quality STEM instruction (21.68%). Notably, many respondents mentioned testing as a negative force in curricular and time constraints and felt that much of their instruction was geared towards state testing.

Leadership

Due to the fact that the leadership aspect of the *Contexts for Teachers' Learning* framework focused heavily on aspects of principal leadership, the quantitative survey measures in the present study did not examine principal leadership in sufficient detail, therefore, predictors of principal leadership were not included in the present study as our sample included only middle grades teachers. However, one theme that emerged from qualitative analysis of open-ended responses was a *lack of administrative support*, with just under 5% of responses (4.90%) in this category.

Time and Funding

Though time and funding could be seen as non-malleable and non-alterable in many cases, as school budgets are largely determined by local tax bases and other factors beyond the control of the people who work in K-12 schools. However, judicious and/or creative use of both time and funding can drastically contribute impact the success or failure of improvement in students' STEM outcomes. Two predictors were related to the allocation of *time and funding,* including teachers' perceptions on the availability of equipment and supplies (ResourceAdequacy) and availability of different types of instructional technology (TechAvailability). As teachers' perceptions of the availability of instructional technology increased, their use of reform-based instructional practices increased by 1.92 points on average. In addition, the largest category of responses regarding instructional barriers were classified into the time and funding category, with six emergent themes classified in this larger category. The six themes, comprising 45.45% of coded responses, included instructional time, funding, resources, technology for students and teachers, a lack of space, and overcrowding of classes due to too many students.

Student/Teacher/Parent Relationships

One aspect not addressed by the *Contexts for Teachers' Learning* framework, relationships among students, parents, and teachers, emerged as a perceived barrier to high quality STEM instruction in just over 15% of coded responses (15.38%). Four emergent themes addressed relationships and/or a need to cultivate stronger relationships in order to address behavioral and/or academic readiness issues. The four themes included *a lack of student engagement or motivation, student academic readiness, student behavior,* and *family or parental support.* In addition to perceiving various aspects of relationships as a barrier, 5% of teachers also perceived relationships through engagement or classroom management as a STEM pedagogical strategy. Responses such as, "good behavior and administrative support for good behavior is important to teaching and learning for everyone" and "have fun and have the kids earn your trust," contrasted with statements such as, "strong classroom management skills, having students learn social skills, and recognize how to make better decisions to decrease their attendance in ISS (In-school suspension)." However, all statements highlighted the importance of student-teacher relationships for high quality instruction.

This study examined the degree to which three aspects of the *Contexts for Teachers' Learning* framework (National Academies of Sciences, Engineering, and Medicine's Committee on Strengthening Science Education, 2015) appeared to impact Texas middle grades STEM teachers' instructional objectives and practices. Statistically significant predictors were found in the areas of building teachers' *professional capacity* and the provision of adequate *time and/or funding*. Regarding professional capacity, as teachers' perceptions of their level of instructional preparedness (*InstrPreparedness*) increased, their use of reform-based objectives and practices increased significantly. In addition, as teachers' perceptions of the availability of instructional technology increased, their use of reform-based instructional practices increased by 1.92 points on average.

This study's results also show that the use of reform-based instructional objectives may have differential impacts by non-malleable factors, including years of experience and content area. Statistically significant control predictors for the use of reform-based instructional objectives included years of teaching experience and primary subject area. Teachers with 10 – 14 years of teaching experience had a reform-based instructional objectives score that was 0.60 points higher, on average, than teachers with fewer than five years of experience. Finally, mathematics teachers had a reform-based instructional objectives score that was 0.58 points lower, on average, than science teachers. A recent study of the professional development of early career teachers (Gabriel, 2010) posited that differentiated professional development could provide teachers' with opportunities to focus their professional learning in ways that would help them grow more efficiently as instructors and may contribute to teacher retention in the field. None of the control variables, including whether or not a teacher taught at a Gold Ribbon campus, significantly predicted teachers' use of reform-based instructional practices.

The emergent themes from analysis of teachers' responses to the two open-ended survey questions overlapped quantitative findings in several ways. Perhaps most notably, the largest category of responses regarding instructional barriers were classified into the time and funding category, with 45.45% of coded responses expressing barriers in this category. This corresponds with the statistically significant and positive impact that teachers' perceptions of the availability of instructional technology had on their use of reform-based instructional practices. In addition, there was quantitative and qualitative triangulation related to developing teachers' professional capacity. In the multiple regression analysis for the use of reform-based instructional objectives, teachers' perceptions of their level of instructional preparedness (*InstrPreparedness*) positively and significantly predicted their use of reform-based objectives. Similarly, qualitative analysis of teachers' perceptions of barriers to STEM instruction showed that teachers also perceived that a *lack of time for collaboration*, *external expectations*, and *teacher motivation* were barriers to high-quality STEM instruction related to developing professional capacity.

Limitations of Study and Future Research

Due to an expanded focus on both fixed and malleable factors, the present study utilized control predictors, such as prior achievement level of students and a school's

Gold Ribbon status, to account for variance explained by non-malleable factors. However, a limitation of the study is its small sample size. Despite multiple reminders, less than 20% of the teachers sampled responded to the survey, rendering multi-level modeling inappropriate as a statistical analysis technique. Therefore, it was not possible to account for the nesting of teachers within schools. In addition, though leadership and teacher/student/parent relationships emerged in qualitative analysis of teachers' perceptions of barriers to high quality STEM instruction, these themes were not examined quantitatively. Finally, due to the manner in which survey questions were routed, non-core STEM teachers, including career and technology and engineering teachers, did not receive all of the questions. Therefore, the listwise deletion used in multiple regression resulted in non-core STEM teachers excluded from analysis of quantitative measures.

The results of this study present several opportunities for future research. Triangulation of quantitative and qualitative data showed that building teachers' professional capacity through instructional preparedness, opportunities for collaboration, an examination of the degree to which external expectations help or hinder teacher growth, and building teacher motivation are key areas of focus for encouraging teacher use of reform-based instructional objectives and practices. Further study of how these areas interact to support teacher change within specific schools or districts is critical to understanding how to best support building teachers' professional capacity. In addition, though both leadership and relationships were present in participants' open-ended responses, further and large-scale study of these two aspects of STEM teaching is necessary, as most studies of leadership and home-school involvement are not STEM-specific. A recent study of a school system-wide program targeted at building stronger home-school relationships found statistically significant increases in mathematics and English/language arts achievement, classroom behavior, and parent involvement for students receiving teacher home visits (Wright, Shields, Black, & Waxman, 2018). However, though the study included STEM subjects and was conducted using data from a charter school system with an explicit STEM focus, the study did not include data on how, or if, building relationships is connected to changes in STEM instructional practice.

Conclusion

The present study provides an in-depth look at STEM teaching practices in a sample of middle grades STEM teachers in Texas. Triangulation of quantitative and qualitative data showed that building teachers' professional capacity through instructional preparedness, opportunities for collaboration, an examination of the degree to which external expectations help or hinder teacher growth, and building teacher motivation are key areas of focus for encouraging teacher use of reform-based instructional objectives and practices.

CHAPTER V

SUMMARY AND CONCLUSIONS

The overarching purpose of this set of studies was to examine both research and practice related to middle grades students' STEM-related achievement, activities, access, and attitudes. Three main questions summarize the focus of this dissertation:

(1) What practices does research identify as most effective for middle grades students in STEM?

(2) To what extent are teachers utilizing effective practices?

(3) What school and teacher factors help or hinder the use of effective STEM practices?

The three studies provide an up-to-date analysis of both the effectiveness of STEM instructional practices in the middle grades as well as the degree to which effective practices identified in the STEM education research literature are being implemented in middle grades classrooms nationwide, as well as in the state of Texas. The studies may also provide information on the degree to which school and classroom contextual factors are promoting or hindering teachers' use of effective STEM practices, as well as the degree to which teachers' use of effective STEM practices may or may not impact student achievement in STEM across the state of Texas.

The meta-analysis conducted in the first study examined the impact of STEM interventions on middle grades students' achievement and attitudes in STEM subjects, as well as what factors moderate the impact of STEM interventions. On average, students involved in STEM interventions performed 0.424 standard deviations higher than students in the control group in experimental studies or prior to the intervention in pre/post studies. The overall average effect was statistically significant ($p < .001$), with a slightly higher effect for achievement (Hedge's $g = 0.608$, $p < .001$). In contrast, the average across all attitude measures was small, but still statistically significant (Hedge's $g = 0.096$, $p = 004$), with students' STEM attitudes 0.096 standard deviations higher than non-intervention students or prior to an intervention. The study's results with other recent reviews of the impact of STEM programs on similar constructs that also found STEM-focused interventions to have small to moderate significant effects on students (An, 2013; Schroeder, Scott, Tolson, Huang, & Lee, 2007). Significant moderators included grade level, study duration, test type, intervention type, and student gender.

The second study utilized hierarchical linear modeling and multiple linear regression to examine which aspects of the *Contexts for Teachers' Learning* framework (National Academies of Sciences, Engineering, and Medicine's Committee on Strengthening Science Education, 2015), includinging professional capacity, coherent instructional guidance, leadership, and providing adequate time and funding; impact middle grades stem teachers' instructional objectives, practices, and instructional technology use. Statistically significant school- and teacher-level predictors in all four areas of the framework suggest building teachers' professional capacity, coherent instructional guidance, leadership, and providing adequate time and funding are impactful areas of study for those seeking to examine stem instructional reform in the middle grades.

The final study took an in-depth look at STEM teaching practices in a sample of middle grades STEM teachers in Texas. Survey analysis focused on teachers' perceptions of reform-based instructional objectives, instructional practices, and use of instructional technology. The main purpose of analysis was to examine the impact of malleable factors for STEM improvement drawn from the *Contexts for Teachers' Learning* framework (National Academies of Sciences, Engineering, and Medicine's Committee on Strengthening Science Education, 2015). Statistically significant predictors were found in the areas of building teachers' *professional capacity* and the provision of adequate *time and/or funding*. Triangulation of quantitative and qualitative data showed that building teachers' professional capacity through instructional preparedness, opportunities for collaboration, an examination of the degree to which external expectations help or hinder teacher growth, and building teacher motivation are key areas of focus for encouraging teacher use of reform-based instructional objectives and practices.

Conclusion

There is no shortage of recommendations for improving middle grades STEM education. However, a great deal of current and prior research centers around student and teacher factors, such as students' socio-economic status and years of teaching experience, that are not alterable or actionable at the school and classroom levels. Though the importance of students' demographic factors and teacher characteristics cannot be ignored, the focus of the present set of studies on a core set of malleable factors at the school and classroom levels has the potential to provide actionable findings to education policy makers, and most importantly, middle grades district and school personnel. The study findings may help to facilitate school- and classroom-level changes to both school environment and instruction that may result in schools structuring teaching environments and instruction in ways that previous research has shown to encourage more underrepresented students to enter and persist through the STEM pipeline.

REFERENCES

Allensworth, E., & Hart, H. (2018). *How do principals influence student achievement?* Chicago, IL: University of Chicago Consortium on School Research. Retrieved from https://consortium.uchicago.edu/sites/default/files/publications/Leadership%20Snapshot-Mar2018-Consortium.pdf

Allensworth, E., Ponisciak, S., & Mazzeo, C. (2009, June). The schools teachers leave: Teacher mobility in Chicago Public Schools. Consortium on Chicago School Research at the University of Chicago Urban Education Institute. Retrieved from https://consortium.uchicago.edu/sites/default/files/publications/CCSR_Teacher_Mobility.pdf

American Statistical Association. (2007). Guidelines for assessment and instruction in statistics education (GAISE) Report: A preK-12 curriculum framework. Retrieved from http://www.amstat.org/asa/files/pdfs/GAISE/GAISEPreK-12_Full.pdf

An, D. (2013). A meta-analysis of the effectiveness of STEM-programs in the United States. ETD Archive. Paper 75. Retrieved from https://engagedscholarship.csuohio.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1074&context=etdarchive

Anzures-Cabrera, J. & Higgins, J. P. (2010). Graphical displays for meta-analysis: An overview with suggestions for practice. *Research Synthesis Methods, 1*, 66-80.

Aschbacher, P. R., Ing, M., & Tsai, S. M. (2014). Is science me? Exploring middle school students' STE-M career aspirations. *Journal of Science Education and Technology, 23*, 735-743.

Aschbacher, P. R., Li, E., & Roth, E. J. (2010). Is science me? High school students' identities, participation and aspirations in science, engineering, and medicine. *Journal of Research in Science Teaching, 47*(5), 564-582.

Bailey, A., Kaufman, E., & Subotic, S. (2015). Education, technology, and the 21st century skills gap. Retrieved from https://www.bcgperspectives.com/content/articles/public_sector_education_technology_twenty_first_century_skills_gap_wef/

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*(1), 1-26.

Beede, D., Julian, T., Khan, B., Lehrman, R., McKittrick, G., Langdon, D., & Doms, M. (2011). Education supports racial and ethnic equality in STEM (ESA Issue Brief No. 05–11). Washington, DC: U.S. Department of Commerce. Retrieved November 10, 2015, from http://www.esa.doc.gov/sites/default/files/reports/documents/educationsupportsracialandethnicequalityinstem_0.pdf.

Biesinger, K. D., Crippen, K. J., & Muis, K. R. (2008). The impact of block scheduling on student motivation and classroom practice in mathematics. *NASSP Bulletin*, *92*(3), 191-208.

Berryman, S. E. (1983). *Who will do science? Trends and their causes in minority and female representation among holders of advanced degrees in science and mathematics. A special report*. New York: Rockefeller Foundation. Retrieved from https://files.eric.ed.gov/fulltext/ED245052.pdf

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. (2009). *Introduction to meta-analysis*. West Sussex, UK: John Wiley & Sons, Ltd.

Brophy, S., Klein, S., Portsmore, M., & Rogers, C. (2008). Advancing engineering education in P-12 classrooms. *Journal of Engineering Education, 97*(3), 369-387.

Brotman, J. S., & Moore, F. M. (2008). Girls and science: A review of four themes in the science education literature. *Journal of Research in Science Teaching, 45*(9), 971-1002.

Brown, P. L., Concannon, J. P., Marx, D., Donaldson, C. W., & Black, A. (2016). An examination of middle grades students' STEM self-efficacy with relation to interest and perceptions of STEM. *Journal of STEM Education, 17*(3), 38.

Bryk, A., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York: Russell Sage Foundation.

Bryk, A., Sebring, P., Allensworth, E., Suppescu, S., & Easton, J. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago, IL: The University of Chicago Press.

Carnegie Commission for Mathematics and Science Teaching (2009). *The opportunity equation: Transforming mathematics and science education for citizenship and the global economy*. New York: Carnegie Corporation of New York and the Institute of Advanced Study, Princeton University.

Catsambis, S. (1995). Gender, race, ethnicity, and science education in the middle grades. *Journal of Research in Science Teaching, 32*(3), 243-257.

Change the Equation. (2016, May). *Left to chance: U. S. middle schoolers lack in-depth experience with technology and engineering.* Vital Signs: Reports on the Condition of STEM Learning in the U. S. Washington, DC: Change the Equation. Retrieved from https://www.ecs.org/wp-content/uploads/TEL-Report_0.pdf

Cheung, A. C., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review, 9,* 88-113.

Children At Risk. (2017, June). The state of public education in Texas: 2017 school rankings analysis. Retrieved from http://childrenatrisk.org/wp-content/uploads/2017/08/C@R-School-Rankings-Webinar-VF1.pdf

Christensen, R., Knezek, G., & Tyler-Wood, T. (2015). Alignment of hands-on stem engagement activities with positive STEM dispositions in secondary school students. *Journal of Science Education and Technology, 24*, 898-909.

Chouinard, R., Karsenti, T., & Roy, N. (2007). Relations among competence beliefs, utility value, achievement goals, and effort in mathematics. *British Journal of Educational Psychology, 77*, 501-517.

Coble, C. (2012). *Developing the analytic framework: Assessing innovation and quality design in science and mathematics teacher preparation*. Washington, DC: Association of Public and Land-Grant Universities (APLU). Retrieved from http://www.aplu.org/projects-and-initiatives/stem-education/SMTI_Library/developing-the-analytic-framework-a-tool-for-supporting-innovation-and-quality-design-in-the-preparation-and-development-of-science-and-mathematics-teachers

Cohen, D.K. and Hill, H.C. (2000). Instructional policy and classroom performance: The mathematics reform in California. *Teachers College Record, 102*(2), 294-343.

Cohen, D.K., and Hill, H.C. (2001). *Learning Policy*. New Haven, CT: Yale University Press.

Cohen, D.K., Raudenbush, S.W., and Ball, D.L. (2003). Resources, instruction and research. *Educational Evaluation and Policy Analysis, 25*(2), 1-24.

Colvin, W., Lyden, S., & Leon de la Barra, B. A. (2013). Attracting girls to civil engineering through hands-on activities that reveal the communal goals and values of the profession. *Leadership and Management in Engineering, 13*(1), 35-41.

Committee on Underrepresented Groups and the Expansion of the Science and Engineering Workforce Pipeline; Committee on Science, Engineering, and Public Policy; Policy and Global Affairs; National Academy of Sciences. (2011). Expanding underrepresented minority participation: America's science and technology talent at the crossroads. Washington, DC: National Academies Press.

Comprehensive Meta-Analysis. (2014). Comprehensive Meta-Analysis (Version 2) [Computer software]. Englewood, NJ: Biostat.

Council of Chief State School Officers' Interstate Teacher Assessment and Support Consortium (2011, April). InTASC model core teaching standards: A resource for state dialogue. Retrieved from http://www.ccsso.org/Documents/2011/InTASC_Model_Core_Teaching_Standards_2011.pdf

Craig, C. J., Verma, R., Stokes, D., Evans, P., & Abrol, B. (2018). The influence of parents on undergraduate and graduate students' entering the STEM disciplines and STEM careers. *International Journal of Science Education, 40*(6), 621-643.

Dabney, K. P., Tai, R. H., Almarode, J. T., Miller-Friedmann, J. H., Sonnert, G., Sadler, P. M., & Hazari, Z. (2012) Out-of-school time science activities and their association with career interest in STEM. *International Journal of Science Education, 2*(1), 63-79.

Dare, E. A., & Roehrig, G. H. (2016). "If I had to do it, then I would": Understanding early middle school students' perceptions of physics and physics-related careers by gender. *Physical Review Physics Education Research, 12*, 1-11.

DeAngelis, K. J., & Presley, J. B. (2011). Teacher qualifications and school climate: Examining their interrelationship for school improvement. *Leadership and Policy in Schools, 10*, 84-120.

Duval, S. and Tweedie, R. (2000), Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*, 455–463.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*, 109-32.

Enders, C. (2010). *Applied missing data analysis.* New York, NY: Guilford Press.

Federici, R. A., Skaalvik, E. M., & Tangen, T. N. (2015). Students' perceptions of the goal structure in mathematics classrooms: relations with goal orientations, mathematics anxiety, and help-seeking behavior. *International Education Studies, 8*(3), 146-158.

Gabriel, R. (2010). The case for differentiated professional support: Toward a phase theory of teacher development. *Journal of Curriculum and Instruction, 4*(1), 86-95.

Garet, M., Wayne, A., Stancavage, F., Taylor, J., Walters, K., Song, M., . . . & Hurlburt, S. (2010). *Middle School Mathematics Professional Development Impact Study: Findings After the First Year of Implementation.* NCEE 2010-4009), U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Washington, DC. Retrieved from https://ies.ed.gov/ncee/pubs/20104009/pdf/20104010.pdf

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine De Gruyter.

Hanson, S. (1996). *Lost talent: Women in the sciences*. Philadelphia: Temple University Press.

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.

Hattie, J. (2009). *Visible Learning*. Abingdon, Oxon: Routledge.

Hill, C., Corbett, C., & St. Rose, A. (2010). Why so few? Women in science, technology, engineering, and mathematics. Washington, DC: American Association of University Women.

Hinojosa, T., Rapaport, A., Jaciw, A., LiCalsi, C., & Zacamy, J. (2016). Exploring the foundations of the future STEM workforce: K–12 indicators of postsecondary STEM success (REL 2016–122). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from http://ies.ed.gov/ncee/edlabs.

Ing, M., & Nylund-Gibson, K. (2017). The importance of early attitudes toward mathematics and science. *Teachers College Record, 119*, 1-32.

Institute of Education Sciences, U. S. Department of Education & the National Science Foundation. (2013, August). Common guidelines for education research and development. Retrieved from https://ies.ed.gov/pdf/CommonGuidelines.pdf

Kim, H. (2016). Inquiry-based science and technology enrichment program for middle school-aged female students. *Journal of Science Education Technology, 25*, 174-186.

Lane, F. C., To, Y. M., Shelley, K., & Henson, R. K. (2012). An illustrative example of propensity score matching with education research. *Career and Technical Education Research, 37*(3), 187-212.

Lara-Alecio, R., Tong, F., Irby, B. J., Guerrero, C., Huerta, M., & Fan, Y. (2012). The effect of an instructional intervention on middle school English learners' science and English reading achievement. *Journal of Research in Science Teaching, 49*(8), 987-1011.

Le, V. Stecher, B. M., Lockwood, J. R., Hamilton, L. S., Robyn, A., Williams, V. L., Ryan, G., . . ., Klein, S. P. (2006). Improving mathematics and science Education: A longitudinal investigation of the relationship between reform-oriented instruction and student achievement [Monograph]. Santa Monica, CA: RAND Corporation. Retreived from https://www.rand.org/pubs/monographs/MG480.html. Also available in print form.

Lipsey, M. A., & Wilson, D. B. (2000). *Practical meta-analysis.* Newbury Park, CA: SAGE.

Llosa, L., Lee, O., Jiang, F., Haas, A., O'Connor, C., Van Booven, C. D., & Kieffer, M. J. (2016). Impact of a large-scale science intervention focused on English language learners. *American Educational Research Journal, 53*(2), 395-424.

Loewenthal, K. M. (2001). *An introduction to psychological tests and scales* (2nd ed.). Philadelphia, PA: Psychology Press.

Loucks-Horsley, S., K. E. Stiles, S. Mundry, N. Love, and P. W. Hewson. (2010). *Designing professional development for teachers of mathematics and science*. Thousand Oaks, CA: Corwin.

Maltese, A. V., & Tai, R. H. (2011). Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among U. S. students. *Science Education Policy, 95*(5), 877-907.

McFarland, J., Hussar, B., de Brey, C., Snyder, T., Wang, X., Wilkinson-Flicker, S., Gebrekristos, S… Hinz, S. (2017). The Condition of Education 2017 (NCES 2017-144). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2017144.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PloS Med* 6(7): e1000097.

National Academy of Sciences, National Academy of Engineering, and Institute of Medicine. (2007). *Rising above the gathering storm: Energizing and employing America for a brighter economic future*. Washington, DC: The National Academies Press.

National Academy of Sciences, National Academy of Engineering, and Institute of Medicine. (2010). *Rising above the gathering storm, revisited: Rapidly approaching category 5*. Washington, DC: The National Academies Press.

National Academy of Engineering; Institute of Medicine (2011). *Expanding underrepresented minority participation: America's science and technology talent at the crossroads*. Washington, DC: National Academies Press.

National Academies of Sciences, Engineering, and Medicine. 2015. *Science teachers' learning: Enhancing opportunities, creating supportive contexts*. Washington, DC: The National Academies Press. https://doi.org/10.17226/21836.

National Action Council for Minorities in Engineering. (2013). *Underrepresented minorities in STEM.* Retrieved from http://www.nacme.org/underrepresented-minorities

National Board for Professional Teaching Standards. (2016). *What teachers should know and be able to do* (2nd Ed.). Retrieved from http://accomplishedteacher.org/wp-content/uploads/2016/12/NBPTS-What-Teachers-Should-Know-and-Be-Able-to-Do-.pdf

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform: A report to the Nation and the Secretary of Education, United States Department of Education*. Washington, D.C.: The National Commission on Excellence in Education.

National Governors Association, the Council of Chief State School Officers, & Achieve, Inc. (2008). Benchmarking for success: Ensuring U.S. students receive a world-class education. Retrieved November 28, 2014 from http://www.edweek.org/media/benchmakring%20for%20success%20dec%202008%20final.pdf

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Washington, DC: The National Academies Press.

Nugent, G., Barker, B., Welch, G., Grandgenett, N., Wu, C., & Nelson, C. (2015). A model of factors contributing to stem learning and career orientation. *International Journal of Science Education, 37*(7), 1067-1088.

Oakes, J. (1990). Opportunities, achievement, and choice: Women and minority students in science and mathematics. *Review of Research in Education, 16*, 153–222.

Ouzzani, M., Hammady, H., Fedorowicz, Z, & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews, 5*, 2-10.

Presley, J., & Coble, C. (2012). *Seeking consensus on the essential attributes of quality mathematics and science teacher preparation programs*. APLU/SMTI, Paper No. 6. Washington, DC: Association of Public and Land-Grant Universities (APLU).

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

Riegle-Crumb, C., Moore, C., & Ramos-Wada, A. (2010). Who wants to have a career in science or math? Exploring adolescents' future aspirations by gender and race/ethnicity. *Science Education, 95*(3), 458-476.

Riegle-Crumb, C., King, B., Grodsky, E., & Muller, C. (2011). The more things change, the more they stay the same? Prior achievement fails to explain gender inequality in entry into stem college majors over time. *American Education Research Journal, 49*(6), 1048-1073.

Rivkin, S. G., Hanushek, E. A., and Kain , J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417-458.

Rothwell, J. (2013, June). *The hidden STEM economy*. Washington, DC: Brookings. Retrieved from https://www.brookings.edu/research/the-hidden-stem-economy/

Rowan, B., Corenti, R., Miller, R., and Camburn, E. (2009). School improvement by design: Lessons from a study of comprehensive school reform programs. In G. Sykes, B. Schneider, and D. Plan (Eds.), *Handbook of Education Policy Research*, Washington, DC: American Educational Research Association. Available: http://www.cpre.org/ school-improvement-design-lessons-study-comprehensive-school-reform-programs [June 2015].

Sanborn, R., Canales, A., Everitt, S., McClendon, D., McConnell, K., O'Quinn, K., & Treacy, C. (2017, May). Children At Risk: 2017 Texas public school rankings methodology. Retrieved from http://childrenatrisk.org/2017-school-rankings

Sass, T. R. (2015). Understanding the STEM pipeline. National Center for Analysis of Longitudinal Data in Education Research (Working Paper No. 125). Retrieved from https://caldercenter.org/sites/default/files/WP%20125.pdf

Schroeder, C. M., Scott, T. P., Tolson, H., Huang, T. Y., & Lee, Y. H. (2007). A meta-analysis of national research: Effects of teaching strategies on student achievement in science in the United States. *Journal for Research in Science Teaching, 44*(10), 1436-1460.

Snyder, T.D., de Brey, C., & Dillow, S.A. (2016). *Digest of Education Statistics 2015 (NCES 2016-014)*. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Retrieved from https://nces.ed.gov/pubs2016/2016014.pdf

Tashakkori, A., & Teddlie, C. (2003). Handbook of mixed methods in social and behavioral research. Thousand Oaks, CA: Sage.

Thavaneswaran, A., & Lix, L. (2008). Propensity score matching in observational studies. Retrieved from https://umanitoba.ca/faculties/health_sciences/medicine/units/chs/departmental_units/mchp/protocol/media/propensity_score_matching.pdf

Texas Workforce Investment Council. (2015, December). Defining middle-skill science, technology, engineering, and mathematics (STEM) occupations in Texas. Retrieved from https://gov.texas.gov/uploads/files/organization/twic/Middle-Skill_STEM_Occupations_in_TX.pdf

U. S. Bureau of Labor Statistics, Employment Projections Program. (2017). Employment projections: Fastest growing occupations. Retrieved from https://www.bls.gov/emp/ep_table_103.htm

Weis, A. M., & Banilower, E. R. (2014, April). 2012 National survey of science and mathematics education: Public release datasets user manual. Chapel Hill, NC: Horizon Research, Inc.

Wilson, S. M. (2011, May). Effective STEM teacher induction, preparation, and professional development. Paper presented at the Workshop on Successful STEM Education in K-12 Schools, Board on Science Education, Washington, D.C.

Wilson, S. M. (2013). Recent developments in STEM Education relevant to the qualities of teacher preparation programs. Storrs, CN: University of Connecticut.

Wright, K. B., Shields, S. M., Black, K., & Waxman, H. C. (2018). The effects of teacher home visits on student behavior, student academic achievement, and parent involvement. *School Community Journal, 28*(1), 67-90.

APPENDIX A

SEARCH TERMS

(Student achievement OR
Academic achievement OR
Achievement OR
Raising achievement OR
Achievement gap OR
High achieving OR
Low achieving OR
At-risk students OR
High needs students OR
Attitude OR
Behavior)
AND
(African American OR
Latin* OR
Hispanic OR
English language learner OR
Minority OR
Students of color)
AND
(Urban education OR
Title I OR
Socio-economic OR
Economically disadvantaged)
AND
(STEM OR
Science OR
Technology OR
Engineering OR
Mathematics)
AND
(Middle OR
Secondary OR
Grade 6 OR
Grade 7 OR
Grade 8

APPENDIX B

CODE SHEET

Citation Information
1. Study title
2. Citation number
3. Publication year
4. STEM subject: 1 = Science, 2 = Technology, 3 = Engineering, 4 = Mathematics, 5 = Mixed subjects
Research Questions
1. Relevant research questions/hypotheses
2. Relevant independent variables (i.e., inputs)
3. Relevant dependent variables (i.e., outcomes)
Participants
1. Grade level
2. Location: 1 =urban US, 2 = suburban US, 3 = Rural US, 4 = Outside US, 5 = Unknown (not reported)
3. Ability level of participants: 1 = low, 2 = average, 3 = high, 4 = mixed, 5 = unspecified
4. Socio-economic status of particpants: 1 = < 50% low SES, 2 >= 50% low SES, 3 = Unknown
5. Race/Ethnicity: 1 = African American; 2 = Asian; 3 = Hispanic/Latino/a; 4 = Native American; 5 = White; 6 = Other, 7 = Unknown
6. Gender of participants: 1 = > 50% male, 2 = > 50% female, 3 = 50% male and 50% female, 4 = Unknown
7. At risk students: 1 = > 50% at risk, 2 = <= 50% at risk, 3 = Unknown
Study Description
1. Research design: 1 = experimental (intervention), 2 = correlational (no intervention)
2. If experimental: 1 = random assignment with control group, 2 = non-random assignment with control group (quasi-experimental), 3 = No control group
3. If control group, how was group equivalence established: 1 = pretest, 2 = propensity score matching, 3 = other, 4 = Unknown
4. If control group, what did control receive: 1 = nothing, 2 = deferred treatment, 3 = alternative treatment, 4 = compensation, 5 = other
5. Total sample size
6. Control group sample size
7. Treatment group sample size

8. If intervention, duration of intervention: 1 = one day or less; 2 = one week or less; 3 = one month or less; 4 = one marking period or less; 5 = one semester (or summer) or less; 6 = one school year or less; 7 = more than one school year (specify)
9. Type of measures: 1 = pre only, 2 = post only, 3 = pre/post, 4 = multiple measurement points, 5 = single data point (i.e., observation), 6 = unknown
11. If assessment, type of assessment: 1 = researcher-created, 2 = previously validated instrument (i.e., survey), 3 = standardized test
10. Describe the type of intervention. 1 = Student-focused intervention; 2 = Teacher professional development intervention, 3 = School-level intervention, 4 = Other (Briefly describe intervention for all codes)
11. Implementer: 1 = researcher, 2 = teacher, 3 = other
12. Intensity of intervention (# of sessions): Put number if known
13. Intensity of session (# minutes per session): Put minutes if known
14. Format: 1 = whole group 2 = small group, 3 = pairs, 4 = individual students
15. Assignment to intervention: 1 = random assignment per student, 2 = random assignment per class, 3 = no random assignment, 4 = unknown
Threats to Study Validity
1. Maturation (subjects mature over course of treatment): 1 = yes, 2 = no
2. Testing (repeated measures): 1 = yes, 2 = no
3. Instrumentation (observer bias, etc): 1 = yes, 2 = no
4. Regression (if you select from extreme high/low); 1 = yes, 2 = no
5. History (events outside of intervention); 1 = yes, 2 = no
6. Selection (overlap in groups before intervention): 1 = yes, 2 = no
7. Other (specify)
Effect Size Data
1. Statistical analysis used: 1 = t-test; 2 = correlations; 3 = ANOVA with post-hoc; 4 = ANOVA without post-hoc; 5 = ANCOVA; 6 = regression; 7 = MANOVA, 8 = other (specify)
2. Effect size information: 1 = Cohen's d, 2 = r^2 (coefficient of determination), 3 = Pearson r (regression coefficient), 4 = Glass' delta, 5 = Hedges g, 6 = odds ratio, 7 = eta^2, 8 = partial eta^2, 10 = omega^2, 11 = beta weights, 12 = unspecified type
Statistical Data
1. Sample size for treatment group.
2. Sample size for comparison (control) group.
3. Treatment group pre-mean.
4. Treatment group pre-standard deviation.

5. Treatment group post-mean.
6. Treatment group post-standard deviation.
7. Comparison group pre-mean.
8. Comparison group pre-standard deviation.
9. Comparison group post-mean.
10. Comparison group post-standard deviation.
Other information
1. If no effect sizes or means/SD, what type of raw data was reported? (describe type of data, for what groups, and page numbers)
Notes

APPENDIX C

LIST OF RETAINED STUDIES FOR META-ANALYSIS

Ball, C., Huang, K. T., Cotton, S. R., & Rikard, R. V. (2017). Pressurizing the STEM pipeline: An Expectancy-value Theory analysis of youths' STEM attitudes. *Journal of Science Education and Technology, 26*, 372-382.

Brown, P. L., Concannon, J. P., Marx, D., Donaldson, C. W., & Black, A. (2016). An examination of middle school students' STEM self-efficacy with relation to interest and perceptions of STEM. *Journal of STEM Education, 17*(3), 27-38.

Chen, C. H., & Howard, B. (2010). Effect of live simulation on middle school students' attitudes and learning toward science. *Educational Technology and Society, 13*(1), 133-139.

Choi, N., & Chang, M. (2011). Interplay among school climate, gender, attitude toward mathematics, and mathematics performance of middle school students. *Middle Grades Research Journal, 6*(1), 15-28.

Hayden, K., Ouyang, Y., Scinski, L., Olszewski, B., & Bielfeldt, T. (2011). Increasing student interest and attitudes in STEM: Professional development and activities to engage and inspire learners. *Contemporary Issues in Technology and Teacher Education, 11*(1), 47-69.

Hsieh, P., Cho, Y., Liu, M., & Schallert, D. L. (2008). Examining the interplay between middle school students' achievement goals and self-efficacy in a technology-enhanced environment. *American Secondary Education, 36*(3), 33-50.

Jackson, J. K., & Ash, G. (2012). Science achievement for all: Improving science performance and closing achievement gaps. *Journal of Science Teacher Education, 23,* 723-744.

Johnson, C. C., Kahle, J. B., & Fargo, J. D. (2007). A study of the effect of sustained, whole-school professional development on student achievement in science. *Journal of Research in Science Teaching, 44*(6), 775-786.

Knezek, G., Christensen, R., Tyler-Wood, T., & Periathiruvadi, S. (2013). Impact of environmental power monitoring activities on middle school student perceptions of STEM. *Science Education International, 1*, 98-123.

Lara-Alecio, R., Tong, F., Irby, B. J., Guerrero, C., Huerta, M., & Fan, Y. (2012). The effect of an instructional intervention on middle school English learners' science and English reading achievement. *Journal of Research in Science Teaching, 49*(8), 987-1011.

Lee, O., Deaktor, R., Enders, C., & Lambert, J. (2008). Impact of a multiyear professional development intervention on science achievement of culturally and linguistically diverse elementary students. *Journal of Research in Science Teaching, 45*(6), 726-747.

Liu, M., Horton, L., Olmanson, J., & Toprac, P. (2011). A study of learning and motivation in a new media enriched environment for middle school science. *Educational Technology Research and Development, 59*, 249-265.

Llosa, L., Lee, O., Jiang, F., Haas, A., O'Connor, C., Van Booven, C. D., & Kieffer, M. J. (2016). Impact of a large-scale science intervention focused on English language learners. *American Educational Research Journal, 53*(2), 395-424.

Marshall, J. C., & Alston, D. M. (2014). Effective, sustained inquiry-based instruction promotes higher science proficiency among all groups: A 5-year analysis. *Journal of Science Teacher Education, 25*, 807-821.

Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical linear models. Applications and data analysis methods. Thousand Oaks, CA: Sage.

Scott, T. P., Schroeder, C., Tolson, H., Huang, T. Y., & Williams, O. M. (2014). A longitudinal study of a 5th grade science curriculum based on the 5E model. *Science Educator, 23*(1), 49-55.

Zwiep, S. G., & Straits, W. J. (2013). Inquiry science: The gateway to English language proficiency. *Journal of Science Teacher Education, 24*(8), 1315-1331.