Causal Mediation Mechanisms in Inquiry-based Science Professional Development

By

Sun Young Yoon

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Educational Leadership and Policy Analysis)

at the
University of Wisconsin-Madison

2015

Date of final oral examination: 8/18/2015

The dissertation is approved by the following members of the Final Oral Committee:
Dr. Geoffrey Borman (Chair), Professor, Educational Leadership and Policy Analysis
Dr. John Diamond, Hoefs-Bascom Associate Professor, Educational Leadership and Policy Analysis
Dr. Adam Gamoran, President, William T. Grant Foundation
Dr. Peter Goff, Assistant Professor, Educational Leadership and Policy Analysis
Dr. Carolyn Kelley, Professor, Educational Leadership and Policy Analysis

ABSTRACT

This study contributes to better understand the key mediation mechanisms involved in the relationship among teachers' professional development, reformed practices, and student outcomes. In an attempt to unpack the causal mechanisms of professional development, this dissertation raises two main questions: 1) How do reformed classroom practices mediate the relationship between professional development and student outcomes? and 2) How do different levels of school organizational capacity, including professional community, principal support, and teacher buy-in, moderate the effects of professional development? To examine the research questions, this dissertation uses data from the Los Angeles Unified School District (LAUSD). The intervention was randomly assigned to 80 schools, including 40 treatment and 40 control schools. In light of the research questions, the analytical strategies are threefold. First, this study decomposes the effect of an inquiry-based science professional development into the direct and indirect effects using causal mediation effect analysis. Second, although teachers intend to implement key principles of effective professional development, specific implementation gaps undermined and contradicted these efforts. Therefore, this study defines the implementation gap between what this professional development intended to change and what actually has been accomplished. Given the different measures of implementation, this study examines which schools and teacher characteristics contribute more to the changed instruction. The third analysis examines why the professional development yields negative student outcomes in the treatment schools. The findings of this study show that inquiry-based instruction does not significantly mediate the effect of professional development on student outcomes. The results indicate that the mediation effect of inquiry-based instruction between professional development and student outcomes is positive but not significantly different in the control versus the treatment. However, the direct effect of the professional development on student outcomes is negative and significant. The negative direct effect may emphasize the importance of school environment when implementing new programs, including factors such as how capably teachers and principals manage new programs. Regarding the measures of implementation outcomes, this study focuses on the implementation gap and school organizational capacity that may influence the gap. The findings suggest that teachers benefit from the treatment by increasing the number of any inquiry features in their instruction when they perceive that their principals are supportive. This finding indicates that school leadership plays an important role for teachers in implementing new teaching approaches. With respect to the relationship between inquiry-based instruction and students' science performance in the treatment schools, the results suggest that students taught by teachers with fewer than 3 years of teaching experience obtain lower scores than students taught by more experienced teachers. However, students of novice teachers benefit more from the use of inquiry-based instruction than students of more experienced teachers. The results underscore that implementing new instructional practices is multilayered with school resources and organizational capacity. To reduce the implementation deficit, cohesive support from school organizations and leadership is vital. Therefore, this study provides insights into how professional development influences student outcomes as mediated by teachers' transformed instruction, and also why this effort has yielded unexpected outcomes.

ACKNOWELDGEMENT

I would like to sincerely thank my advisor, Dr. Geoffrey Borman, for his continuous guidance, understanding, and patience. Without his encouragement and inspiration, this work would not have been possible. I would also like to express my sincere and deepest gratitude to my committee members. I feel blessed and honored to work with my committee members Dr. Carolyn Kelley, Dr. John Diamond, Dr. Adam Gamoran, and Dr. Peter Goff.

I would never have been able to finish my dissertation without my family's support. I am grateful that my parents, my brother, sister-in-law, and my lovely nephew have been always there for me. Finally and most importantly, I would like to thank my beautiful mother, Sook Ja Kim. Her unconditional love and prayers have led me to go through this journey. I believe God will give her enough strength to get better soon as he has given me strength to finish this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST of FIGURES

CHAPTER I

INTRODUCTION

This study examines the causal mediation mechanisms in teachers' inquiry-based science professional development to provide a better understanding of teacher professional development. Currently, there is a significant need for high quality science education. A report entitled *Prepare and Inspire: K-12 Education in Science, Technology, Engineering and Math (STEM) Education for America's Future* (President's Council of Advisors on Science and Technology, 2010) found that "the United States will need 1 million more STEM graduates over the next decade than will be produced by our current modes of education" (Gates & Mirkin 2012). In response to this need, the U.S. government has been increasing funding for improvements in education in the STEM fields. For example, in February of 2012, President Barack Obama proposed 80 million dollars in funding aimed toward the training of more than 100,000 math and science teachers (White House, 2012).

Despite nationwide reform efforts and attention to improving students' academic performance, science achievement in the United States still lags behind other countries (Fleischman, Hopstock, Pelczar, & Shelley, 2010). Research that uses representative national data found that only 14% of science lessons provided to students were of high quality (Banilower, Smith, Pasley, & Weiss, 2006; Banilower, Heck, & Weiss, 2007). The urgent need for improving U.S students' science achievement could not be more evident. One of the vital problems for the science education is teachers' lack of skill and appropriate training (Groome, Rankin, & Wheary, 2012).

To combat this problem, research has lately emphasized inquiry-based science teaching and learning as the key to improving teacher practices and student science outcomes. The National Research Council (2000) noted, “For students to understand inquiry and learn to use it in science, their teachers need to be well versed in inquiry and inquiry-based methods” (2000, p. 87). However, implementing new educational reform curriculum, such as inquiry-based science instruction, into school systems is intellectually demanding and challenging work (Spillane, Reiser, & Gomez, 2006). In turn, since the National Science Foundation first funded a hands-on (or inquiry-based) science curriculum in the 1950s, teachers have often failed to understand its reform-oriented principles and to successfully adopt the new science curricula (Penuel, Fishman, Gallagher, Korbak, & Lopez-Prado, 2008). In addition, about 70% of teachers in elementary and middle school reported a need for professional development that trains instructors in the use of inquiry-based teaching strategies (Weiss, Banilower, McMahon, & Smith, 2001).

As a result, science professional development initiatives have attempted to improve teachers’ content knowledge and teaching strategies. It has been widely believed that professional development plays an essential role in improving teachers’ implementation of science instruction. However, evidence shows that even well-designed professional development often fails to improve teaching practices and student outcomes. The important question is thus, what may cause professional development initiatives to fail? The first scenario can be that professional development fails to improve teaching practices and in turn student outcomes. The second scenario can be that professional development yields changes in instruction, yet it causes no impacts on student outcomes.

Identifying causal pathways of professional development is essential to understand how professional development initiatives contribute to both teachers' and students' outcomes.

When implementing professional development, there may be potential variance in treatment effects due to school organizational capacity, such as the professional learning community, principal support, and teacher buy-in. For instance, teachers' professional learning community is a place where invisible and informal professional learning occurs among teachers (Little, 1993). Teacher learning often occurs at the collective level, which may engender heterogeneity in effects. For example, a study by Fulton, Doerr, and Britton (2010) highlights an important role of professional learning community for STEM teachers in synthesizing content knowledge and positively influencing their instruction. However, prior studies have not clearly demonstrated how the school organizational capacity contributes to effects of professional development.

Therefore, the focus of the current study is twofold: first, this study examines the causal mediation effect of teacher practices between professional development and science student outcomes; second, in order to better explain teachers' learning through professional development, this study identifies variability in effects of professional development according to school organizational capacity including professional learning community, principal support, and teacher buy-in. To do so, this study addresses ways in which teachers change their practices based on different targeted implementation outcomes.

Methodologically, research often overlooks important mediating mechanisms, such as how professional development has an impact on student outcomes through teacher practices, and ways in which the school organizational capacity contributes to heterogeneity in effects of professional development. However, when teachers implement reforms in classrooms, the impacts emerge through multiple paths (Sunal & Wright, 2006). A few studies have examined the mediating relationship of teaching practices between professional development and student outcomes. However, none of these studies reveals a causal relationship. The currently proposed study therefore will contribute to the body of literature by examining how a causal mediation effect of teachers' inquiry-based practices may vary by their school organizational capacity.

One of the ways to understand the mechanisms of professional development is investigating its causal pathway. For example, A can influence B directly or indirectly; if the influence is indirect, then M acts as a mediating factor. When A influences B directly, they (A and B) are in a causal relationship, when indirectly (i.e. with M in the middle) in a mediated relationship. The whole system can be described as an intermittently mediated causal relationship. This study therefore uncovers a causal path of professional development and student outcomes as deriving the causal mediation effect model recently developed by Imai, Keele, and Tingley (2010).

In addressing challenges to improving science achievement, low-achieving urban schools are more likely to suffer from the science achievement gap due to their substantial environmental disadvantages (King, Shumow, & Lietz, 2001). Teachers from low-performing urban school systems, in turn, are more likely to encounter instructional challenges; this is often due to students' less privileged backgrounds. Broadly speaking, research has shown that teachers in schools serving minority students tend to have lower efficacy, lower test scores on teacher certification, and are less prepared to teach science than teachers in schools with better conditions (Johnson & Marx, 2009; King et al., 2001).

In light of the particular obstacles science teachers face in low performing schools serving minorities, this study targets the nation's second largest school district, the Los Angeles Unified School District (LAUSD), which is well known for its large minority and low-income student population. This study investigates the impacts of a large-scale randomized trial of inquiry-based science professional development for elementary school teachers in the LAUSD. This system-wide change intervention is supported by System-Wide Change for All Learners and Educators (SCALE) of the National Science Foundation Math and Science Partnership. Curriculum units of teacher professional development include inquiry-based science and are known as "immersion units."

In an attempt to unpack the causal mechanisms that may identify the ways in which professional development influences student outcomes, this study raises two main questions.

1) How do reformed classroom practices mediate the relationship between professional development and student outcomes?

2) How do different levels of school organizational capacity, including professional community, principal support, and teacher buy-in, moderate the effects of professional development?

CHAPTER II

RESEARCH FRAMEWORK

II. 1. The intervention: Inquiry-based Science Professional Development

The growing need for scientific inquiry-based instruction has been recognized among science educators. While traditional science instruction mainly focuses on content knowledge, inquiry-based science instruction aims to improve students' ability to investigate scientific questions and use evidence to acquire scientific knowledge (Grigg, Kelly, Gamoran, & Borman, 2013). The inquiry-based science instruction intends to better understand science and strengthen students' engagement in the real-scientific world. According to the National Research Council (1996),

> Scientific inquiry refers to the diverse ways in which scientists study the natural world and propose explanations based on the evidence derived from their work. Inquiry also refers to the activities of students in which they develop knowledge and understanding of scientific ideas, as well as an understanding of how scientists study the natural world (NRC 1996, p. 23).

According to the National Science Teachers Association's recommendation (2004), an inquiry-based science education helps students develop content knowledge by guiding them in how to ask questions and to use evidence for answering questions. The NRC presented the five "essential features of classroom inquiry" (NRC, 2000) at the time of the study[1]. The central concepts of the inquiry support students' ability to integrate scientific knowledge and their understanding. The essential feature of inquiry established by the NRC (2000) consists of 5 features of scientific inquiry and variations as shown in Table 1. These 5 features of inquiry lead students to "do" scientific experience as a cycle. These inquiries aim to foster students' cognitive and scientific development (Jeanpierre, Oberhauser, and Freeman, 2005; NRC, 2000).

Feature 1: Learner engages in scientifically oriented questions

Feature 2: Learner gives priority to evidence in responding to questions

Feature 3: Learner formulates explanations from evidence

Feature 4: Learner connects explanations to scientific knowledge

Feature 5: Learner communicates and justifies explanations

[Table 1 Here]

The System-Wide Change for All Learners and Educators (SCALE) partnership developed an inquiry-based science immersion professional development, known as "immersion units" with the University of Wisconsin-Madison in partnership with the University of Pittsburgh and LAUSD. The intervention, a 6 weeks *Rot it Right* science

[1] The National Research Council (2012) has recently recommended the K–12 science standards, highlighting eight scientific and engineering practices:

1. Asking questions (for science) and defining problems (for engineering)
2. Developing and using models
3. Planning and carrying out investigations
4. Analyzing and interpreting data
5. Using mathematics and computational thinking
6. Constructing explanations (for science) and designing solutions (for engineering)
7. Engaging in argument from evidence
8. Obtaining, evaluating, and communicating information

immersion curriculum as the life science standard of fourth grade students, was designed to help teachers engage students in scientific thinking, which includes a full cycle of inquiry features in scientific investigation. The Rot it Right curriculum includes thirteen 45-minutes lessons for 6 weeks, about 10 hours of instructional time. The contents of Rot it Right are: 1) based on the five essential features of inquiry recommended by NRC (2000), 2) key concepts in the California State Science Content Standards for 4th grade life science, and 3) intended to improve students' understanding of science by practicing scientific thinking (Lal & Osisioma, 2010).

This study uses the first-year data from a school-based large-scale randomized field trial of the science immersion professional development; this trial was done with support from the National Science Foundation from 2006 to 2009. The first year of the immersion unit curriculum, called "Rot it Right," targets fourth grade instruction that covers life science standards on basic concepts of energy and matter through food chain and living/ nonliving components of ecosystems and their roles in ecosystem.

II. 2. Contexts of the Study

This study is built on the effort from prior studies that have investigated the effects of the immersion unit. A study that used the same data as the current study investigates a causal impact of professional development on student outcomes (Borman, Gamoran, & Bowdon, 2008). Contrary to what researchers expected to find, findings of the study indicated that inquiry-based science professional development had a negative impact on fourth-grade student outcomes. The treatment schools achieved about one quarter of a standard deviation lower than the comparison schools on district-wide standardized assessments. More importantly, students taught by science lead teachers—mainly targeted for the immersion institutes—achieved approximately one half of a standard deviation lower than the comparison schools. As supplementary results, the authors also found a positive interaction effect for students of novice teachers (3 years experience or less) but a larger, negative effect for students of veteran teachers. This finding corroborates a finding from a study by Clotfelter, Ladd, and Vigdor (2007) that teachers' achievements improve the most in their early careers.

A possible reason for the negative results of professional development is a lack of alignment between reformed instruction and the periodic assessment (Borman et al., 2008). As such, they conducted additional analyses examining whether the treatment group yields better outcomes on the subsets of standards on the life science periodic assessment. The results consistently suggest lower scores in treatment school students compared to students in the control group schools. Another potential reason might stem from teachers' implementation of immersion units.

Consequently, another study examines whether participation in the immersion institutes led to changes in teachers' science instruction (Grigg et al., 2013). They found that scientific inquiry was more frequently observed in the classroom of the treatment schools than the comparison schools. In addition, it should be noted that five essential features of inquiry are intended to be cyclical and non-linear. However, the findings show that teachers failed to implement a full cycle of scientific inquiry into the classroom. For example, feature 1 and 2 are more frequently present than the other three features in the classroom observation (Grigg et al., 2013; Lal & Osisioma, 2010). Therefore, Grigg et al. (2013) concluded that while scientific inquiry yields changes in immersion teachers' instructional use of the way teachers and students interact, it may not transform their scientific understanding.

Though limited in number, research has carefully investigated a link between professional development, teacher practices, and student outcomes, as further indicated in the literature section below. Focusing on a full cycle of scientific inquiry, Jeanpierre et al. (2005) examine effects of science content professional development, funded by NSF. It should be noted that their study relies on teachers' self-reports, while Grigg et al. (2013) and the current study use classroom observation data. The findings show that teachers' participation in the professional development institute significantly supports teachers' implementation of full scientific inquiry into their classrooms.

As another example of investigating the relationship between professional development and student outcomes using large-scale data sets, Wallace (2009) found that professional development occasionally has significant impact on student outcomes on reading and math when teacher practices mediate the effects of professional development. It also indicates that approximately 68% to 85% of the variance in teacher math practices and 50% to 75% of the variation in teachers reading practices were not explained by variables that may affect teaching practices. Yet the method Wallace (2009) uses, the structural equation modeling (SEM) approach, only addresses correlational relations. The currently proposed study therefore will contribute to the literature by examining the causal relationship of professional development using multiple forms of large-scale data.

II.3. Conceptual Framework of the Study

When implementing reform initiatives such as professional development yield unintended results, it is important to understand why the reform initiatives often fail. There is a need for more comprehensive ways to examine effects of professional development. Therefore, this study seeks new evidence for the effects of professional development and extends existing works on causal relations in causal mediation effect models.

The conceptual framework guiding this study draws on theoretical perspectives from research by Gamoran et al. (2000) and Hargreaves and Fullan (2012). These studies capture teacher learning as not only individual teachers' capability but also school capacity to enhance school improvement. This framework has been chosen in order to examine the mechanisms of science instructional reform through professional development, which may include teacher learning, implementing reformed instruction, and community practices.

Integrating the theoretical lenses, this study develops a conceptual framework for understanding how reformed teaching and learning occur through professional development. To emphasize organizational resources, Gamoran et al. (2000) elaborate essential resources, including material, human, social resources: 1) material resources, including curriculum materials, time, equipment, and planning for teaching, focus more on how they were utilized rather than their direct impact on student learning; 2) human resources stress teachers' pedagogical knowledge and teaching skills, which may critically determine quality of teaching and student learning; 3) social resources consist of shared values, collaboration, and collective decision-making.

A more recent study by Hargreaves and Fullan (2012) highlights teachers' professional capital that consists of human, social, and decisional capitals: 1) human capital can be defined as talent, which can be complemented with social capital; 2) social capital plays a role in sustaining individual development; 3) decisional capital can be defined as capacity for judgment with collective responsibility and sharing feedback as teachers encounter uncertainties in their practices. In sum, professional capital is "about collective responsibility not individual autonomy about scientific evidence as well as personal judgment; about being open to one's clients rather than standing on a pedestal above them"(p. xv). While those two studies commonly highlight the importance of human and social capitals, Hargreaves and Fullan (2012) distinctly discuss decisional capital in addition to human and social capital. Clearly, teachers' control over the way they allocate resources is a key determinant of the impact that resources can have (Gamoran et al., 2000).

It should be also noted that those resources are enhanced by being combined with each other. Social capital strengthens teachers' human capital by improving their knowledge (Coleman, 1988; Gamoran et al., 2000; Hargreaves & Fullan, 2012). Human resources can also amplify effects of material resources. For example, class size tends to influence student learning as teachers pay greater attention to individual students and instructional strategies, not because of the number of students in a class itself (Bascia, 2010; Gamoran et al., 2000). These studies reiterate the importance of resources and their influence on the ways in which professional development can help teachers' instructional improvement.

To answer the first research question, this study examines how teaching practices mediate between professional development and student outcomes. This study seeks to understand the total, direct and indirect influence of inquiry-based science professional development on student outcomes. As shown in Figure 1, Teacher Practices (M) is a mediator of the professional development impacts (X) on Student Outcomes (Y). The professional development (X) was randomly assigned at the school level, and the mediator (M) was observed at the school level. Student outcomes (Y) is at the student level, which is nested in the school level. The effect of professional development is mediated through teacher practices on student outcomes. C denotes a direct effect of professional development on student outcomes; A and B indicate indirect influence of professional development on student outcomes.

[Figure 1 Here]

When attempting to change teaching practices, change implies improvement of teaching practices, content knowledge, and instructors' attitudes toward teaching methods (Jeanpierre et al., 2005). Change can sometimes demand collaborative process and does not easily occur because teaching is not a linear behavior that reacts to educational reform policies (Johnson, 2006). It is a rather complex process related to teachers' learning and their perceptions, especially when implementing reformed practices. In light of the complex and collaborative process of implementing instructional reforms, this study considers school organizational capacity as a critical learning resource that helps implementation of professional development. For instance, one of the most frequently cited reasons for the importance of professional learning community is that teachers' collaboration, shared goals and responsibilities, and social trust drive instructional changes in classroom and improve student achievement (Coburn & Stein, 2006; Grodsky & Gamoran, 2003; Little, 2006; Louis & Marks, 1998). Therefore, this study sheds light on how the school organizational capacity as a moderator leads to variation in the effects of inquiry-based science professional development.

CHAPTER III

LITERATURE REVIEW

> “All genuine learning is active, not passive. It involves the use of the mind, not just the memory” (Adler, 1982, p. 50).

> “Teaching isn’t an exact science. Uncertainty is in its nature. This uncertainty calls for wise, well-founded judgment. Uncertainty is the parent of professionalism and the enemy of standardization. It is what makes teaching interesting, variable, and challenging—a job that’s different every day” (Hargreaves & Fullan, 2012, p. 107).

III. 1. Implementation of Inquiry-based Science Professional Development Initiatives and its Link to Instructional Changes

Inquiry-based science instruction has been highlighted as one of the primary factors that improve students’ science achievement. The most central principle is that teachers “who have experience in the practice of science, and in the use of science in the ‘real world,’ can better communicate the concepts and value of science to their students” (Dubner et al., 2001, p.3.6.3). To improve inquiry-based science instruction, professional development has been a vital component. As evidence, the National Science Foundation (NSF) has funded teachers’ professional development for science education for more than 20 years. Given the great amount of investment in teachers’ professional development, policymakers and stakeholders in education increasingly pay attention to evidence on its effects (Ingvarson et al., 2005).

Studies broadly suggest that teacher professional development plays an essential role in how teachers learn and transform acquired knowledge into classroom instruction and in turn improve student achievement (Avalos, 2011; Wallace, 2009). While research has shown positive effects of professional development on instruction, some of the unintended results of professional development support skeptical claims regarding professional development (Gamoran et al., 2000). However, our understanding of the ways in which professional development influences teacher learning and practices remains unclear. To explain possible reasons for those mixed outcomes of professional development, this section particularly focuses on the mediating role of teacher instruction in implementing professional development and the ways teachers learn new reform knowledge and implement them in classroom.

First, even with well-designed professional development, there may exist a gap between what professional development intends to change and what teachers actually implement. Professional development is an important component of reform strategies because it targets changes in teacher practices and student learning (Knapp, 2003). Professional development affects student learning, but only as it is applied by teachers in classroom. In this sense, teachers' instructional improvement demanded by reform efforts primarily count on teachers (Borko, 2004; Fullan & Miles, 1992). However, teachers' acquired knowledge from professional development may not be integrated with reform components, which causes conflicts between their own beliefs about science instruction and actual classroom practices (van Driel, Beijaard, & Verloop, 2001).

For instance, the NSES found that teachers are not likely to successfully implement inquiry-based science instruction since they are not always knowledgeable and often reluctant to implement standards-based instruction (Johnson, 2006). One of the challenges of implementing reform initiatives stem from the different worlds in which teachers live and where educational reforms occur (Lin & Cooney, 2001). In other words, educational reform programs often fail due to the misunderstanding of policy messages and miscommunication between program developers and implementers (Spillane et al., 2006). As Little (1993) also noted, “professional development is where the tension between institutional imperatives and individual prerogatives exists, between the conditions necessary to attempt systemic change and the conditions that engage individual teachers in their work” (p. 141).

Secondly, another possible reason for the gap between what professional development intended to reform and what teachers actually implement may derive from a gap between what professional development intends to change and what teachers actually learned through professional development. Implementation of reform initiatives relies on teachers’ learning (van Driel et al., 2001). However, research often expresses skepticism about whether professional development drives or even equals professional learning (Garet et al., 2001; Groundwater-Smith & Mockler, 2007). Little (2006) also argues that teachers often fail to connect professional development to their own instructional interests and their current instructional problems. As a result, professional development reform initiatives often yield nominal and even unintended outcomes.

It is difficult to change teacher practices rapidly, and even more challenging to determine whether the professional development actually leads to teachers’ meaningful learning (Ingvarson et al., 2005; van Driel et al., 2001). In this sense, Loucks-Horsley et al. (1998) also claim that it is challenging for teachers to teach what they themselves have not learned. It is almost impossible to improve student learning without improving teacher learning (Fullan, 1996). To successfully implement professional development in practice, the materials to be taught need to become a natural part of the teachers' professional skills (Fullan & Miles, 1992).

Research also indicates that increasing teachers' science content knowledge and its implementation through actual experiences can lead to positive changes in their instruction (Jeanpierre et al., 2005). Garet, Porter, Desimone, Birman, and Yoon (2001) asserted that "activities that are content focused, but do not increase teachers' knowledge and skills have a negative association with change in teachers' practice" (p. 934). For example, Newman et al. (2011) found that even if science teachers in the treatment group used reform strategies averaging 40.07 more minutes than control teachers, teacher reported content knowledge was not influenced by the new science learning strategies. These studies show that teachers' professional development should be closely aligned with teachers' professional learning.

However, efforts that investigate the impacts of professional development often overlook its relationship to "teacher learning" (Ball, 1996, Jeanpierre et al., 2005). While most educational reforms have heavily focused on whether a program has an impact on improving student outcomes, we have little understanding about how teachers understand the messages from the reform and actually implement what they learn from them. Therefore, it is important to raise the penetrating and fundamental question whether teachers actually learn from professional development, which in turn leads to changes in their practices (Ingvarson et al., 2005).

Lastly, research has begun to confirm that educational reform efforts fail if they ignore teachers' cognitive process and existing knowledge (van Driel et al., 2001). These gaps in linking professional development and teacher practices underscore the importance of better understanding how teachers shape the sense-making process for meaningful changes of an intended reform (Coburn & Stein, 2006; Henstrand, 2006, Spillane et al., 2006). In other words, it is crucial how well teachers understand the reform principles and adopt them based on their own interpretation. If teachers do not successfully implement reform principles, students' academic progress may not be sustainable. In order for successful and sustainable outcomes, Vygotsky (1978) has theorized that teachers should internalize teaching principles, which highlights the importance of a better understanding of how teachers comprehend knowledge (Kent, 2004). As such, it is of interest to learn how teachers understand reform messages and actually implement them.

As one of the successful ways to implement education reform such as professional development initiatives, researchers emphasize teachers' sense-making process. There have been questions about the ways in which teachers interpret reform based on their prior knowledge and understanding. Research suggests the sense-making process as one of the powerful tools for understanding and adopting policy messages (Coburn, 2001; Coburn & Stein, 2006).

One of the reasons for the importance of teachers' sense-making process is that changes in practices can be more difficult to accomplish if they demand change of teachers' understanding or prior knowledge through active interaction (Johnson, 2006). In particular, inquiry-based science instruction may require teachers to have a higher level of pedagogical content knowledge (Johnson, 2006; Keys & Bryan, 2000). The inquiry-oriented approach to teaching and learning in science education is also known to be a challenge for teachers because it demands changes in classroom management strategies, the organization of knowledge, and assessment (Fishman, Marx, Best & Tal, 2003). Taken together, these studies reiterate the importance of better understanding of a multiple ways in which teachers enact reform principles led by professional development.

III. 2. Inquiry-based Science Teaching Practices in the Classroom and Student Science Achievement

Lack of quality teachers in science is a critical concern in the United States. A large body of research addresses teacher quality and effectiveness as key factors for students' academic improvement. For example, evidence on teacher quality and effectiveness suggests that 21% of the variance in student math achievement is associated with variation in teacher effectiveness (Wallace, 2009). Teacher quality in science is a key determinant that influences students' science achievement and therefore should be improved (Forster, Toma, & Troske, 2013).

Quality instruction demands content knowledge and thus science instruction may require more advanced instructional strategies (Community for Advancing Discovery Research in Education, 2011). Regarding instructional strategies, the National Research Council's NSES (NRC, 1996) suggests key components to enhance quality science instruction. The NSES recommends "teaching for understanding by engaging students with meaningful ideas in science at a conceptual level, rather than focusing on terminology and the ability to recall facts" (Banilower et al., 2006). To facilitate learners' meaningful engagement, teachers should facilitate students' inquiries and help them link the new knowledge to the real world (Banilower et al., 2006).

With respect to students' inquiry-based learning, Ausubel (1967) coined the term "meaningful learning," which is defined as learners' ability to connect new knowledge to existing cognitive structures. Bransford et al. (2006) discuss that students may have difficulties to acquire new knowledge that is taught in the classroom without engaging in significant understanding. Those authors suggest that factual knowledge is a means to build students' understanding rather than an end goal in itself.

However, inquiry-based science instruction is different from how a majority of teachers have been teaching science (Wallace, 2012). This may lead science educators to face the challenge of implementing inquiry-based science instruction, which demands pedagogical practices such as constructivism and project-based learning (Penuel et al., 2008). In light of increasing accountability pressure and demands for broader content knowledge, science professional development is key to implementing inquiry-based science instruction (Fishman et al., 2003; Garet et al., 2001; Groundwater-Smith & Mockler, 2007, Wallace, 2012).

Professional development not only plays a key role in improving teachers' instruction, it also influences student outcomes. Yet, research points out that there is little evidence for effects of professional development on teacher practices as an intermediate outcome and its primary effects on student outcomes (Wallace, 2009). It is critical to understand the distinctive contributions of professional development to instructional improvement and student outcomes (Fishman et al., 2003; Garet et al., 2001; Wallace, 2009).

A rich body of literature has examined how inquiry-based science professional development affects a teacher's science knowledge, teaching strategies, interaction with students, and ultimately students' academic performance (Buczynski & Hansen, 2010; Smith et al., 2007; Supovitz & Turner, 2000). Studies indicate a positive association between inquiry-based science professional development, teacher instruction, and student outcomes. For instance, Akerson and Hanuscin (2007) examined a three-year professional development program that emphasizes scientific inquiry and inquiry-based instruction by tracking science teachers' instructional changes. The interview and observation results show that teachers improved their science pedagogy in the classroom and influenced students' ideas on science by attending professional development programs. Fishman et al. (2003) also found that teachers implemented what they learned from professional development, which in turn improved student outcomes.

On the other hand, recent studies suggest no impacts of professional development on student outcomes (Heller, 2012; Newman et al., 2012,), or even indicate negative impacts (Borman et al., 2008; Martin, Brasiel, Turner, & Wise, 2012). Even if researchers recognize that it is critical to understand the reasons for the unintentional outcomes, it still remains unclear. As Heller (2012) points out, it is hard to determine whether the reasons for the absence of impact come from a failure of implementation or no significant impacts of reformed practices. For example, researchers may overlook whether changes in instruction affect student outcomes in desirable ways. In addition, even if the changes from a new curriculum intend to be beneficial for student outcomes, assessment may not be aligned with the inquiry science reform effort (Borman et al., 2008; Penuel et al., 2008).

To this end, examining the relationship between professional development and student learning through reformed teacher practice remains an important inquiry. Yet, there is little evidence that suggests a causal inference between those relationships (Guskey & Sparks, 2004). This study addresses that gap in the literature by examining how principles from a large-scale professional development program can emerge in the classroom, and in turn analyzing how instructional changes can lead to improvement in student performance.

III. 3. Professional Development and School Organizational Capacity

In this section, the attention is shifted toward school organizational capacity in implementing professional development. Despite a great amount of investment by schools, districts, and federal government, we have a limited understanding of how teachers learn from professional development (Borko, 2004). Most of school reform initiatives, including professional development, are externally designed and adopt "outside-in" models, and may often neglect schools' internal contexts (Coburn, 2001). Therefore, research stresses organizational support for school improvement (Scribner, Cockrell, Cockrell, & Valentine, 1999). In light of the barriers school reformers may face, teachers' mediating roles in both transmitting school culture and reconstructing what they understood are key to improving quality of policy implementation and teacher practices. The relevant question in the current study is, what support system moderates impacts of inquiry-based science professional development? Therefore, this study includes professional learning community, principal support, and teacher buy-in as school organizational capacity to promote effects of professional development.

1) Professional Learning Community

One of the many factors that lead to instructional changes is to facilitate teachers to participate in professional learning community. A professional learning community is an organizational structure that shapes an intellectual culture (Louis & Marks, 1998). The concept of the professional learning community has gained popularity since the 1990s from a perspective of learning organization (Futon & Britton, 2010). Little (2006) defined professional community as "close relationships among teachers as professional colleagues, usually with the implication that these relationships are oriented toward teacher learning and professional development" (p. 15). More recently, Hargreaves and Fullan (2012) define professional learning communities as organizations that facilitate collaborative school improvement based on decisions made within constructive conversations among teachers. The concept of a professional learning community emphasizes learning rather than teaching, and it aims to warrant that students learn as well as to ensure that they are taught (DeFour, 2004).

Research suggests that teacher collaboration, experimentation (risk-taking), and integrating individual goals with organizational goals are key factors in implementing professional development successfully (Groundwater-Smith & Mockler, 2007). When implementing professional development, teachers' accumulated experiences and interaction with their colleagues are critical to decision-making and adopting professional growth into their practices. The changes from implementing reform programs will not occur without teachers' collective effort toward professional learning and growth (Lewis, 2011). With respect to the changes in instruction, decision-making skills are key to implement components from professional development in classrooms. Hargreaves and Fullan (2012) discuss that interaction with other teachers enhances their decisional capital, which in turn leads to more effective practices in their classrooms.

Furthermore, professional learning community plays a critical role in sustaining the effects of professional development (DeFour, 2004). Research highlights the importance of collaborative learning to help teachers successfully implement new curriculum through professional development. A study also suggests that individual teachers gain an advantage not only from professional development but also from participation with their colleagues (Grodsky & Gamoran, 2003). It emphasizes not only individual aspects of teacher learning but also its social aspects (Coburn & Stein, 2006). Increasing evidence highlights the role of the professional learning community in enhancing teachers' knowledge and facilitating teachers' actual learning within professional learning community. Coburn and Stein (2006) also underline "a culture of enquiry within schools" rather than "a culture of compliance" as the best way to achieve a teacher's professional learning through reform initiatives including professional development. Research also suggests that social capital reinforces teachers' human capital (Coleman, 1988; Gamoran et al., 2000; Hargreaves & Fullan, 2012). Human capital can be enhanced by teaching experiences, professional development, and subject knowledge, but it can be also reinforced by strong social capital, such as trust and social networking among colleagues. As Leana (2011)'s study shows, when schools have low human capital but high social capital, they are better situated than schools with high human capital but low social capital. The most finding shows that students obtained higher scores in math achievement when their teachers are more likely to interact with their colleagues (Leana, 2011). In light of a well-established body of literature on the relationship between professional learning community and the impacts of professional development, this study examines whether the level of professional learning community determines impacts of professional development.

2) Principal Support

Principal support is critical for successful implementation of reforms (Datnow & Castellano, 2001). Theoretical lenses shed light on how individual teachers acquire learning by "socially organized interaction" (Greeno et al., 1998, p. 6). With respect to the socially engaged learning in schools, principals play a vital role in shaping the loose coupling between administration and instruction in schools by shaping common goals (Gamoran et al., 2000; Halverson, 2007). Despite a large body of literature on principal leadership, only a few studies have emphasized principal support in implementing professional development for teachers. Therefore, this study focuses on the contribution of principal support in improving effects of professional development.

Research demonstrates that principal leadership indirectly influences student learning by interacting with teachers (Hallinger & Heck, 1996; Hallinger, Bickman, & Davis, 1996). Among a wide range of functions, leadership may provide significant support for teachers' understanding of its elements in implementing school-wide professional development. For example, given the necessary resources for teachers' learning such as a stable and positive learning climate in professional development, principals play a key role in building professional capacity, which links to classroom instruction (Datnow & Castellano, 2001; Sebastian & Allensworth, 2012). With respect to this leadership role, research demonstrates that principal support is considered as the most critical factor in determining the outcome of professional development and building an effective learning environment (Weiss & Pasley, 2006). Strong principal support is intended to influence teachers' learning processes that may mediate the principles from professional development on student outcomes (Sebastian & Allensworth, 2012). However, principal support is often neglected in evaluating the effects of professional development. Therefore, this study empirically examines the influence of principal support for implementation of key components in professional development.

3) Teacher Buy-in

A growing body of research views transformation of teaching practices as a site of policy implementation (Coburn & Stein, 2006). When practicing new instruction, teachers may experience confusion from changes in procedures and pedagogy (Kent, 2004). It is critical that teachers understand the reform principles through professional development and adopt them in meaningful ways, based on their own interpretation. Despite the great effort to improve instruction through professional development, only a few studies focus on teacher buy-in and its relations to outcomes of professional development.

From a cognitive perspective, teachers' implementation process is coupled with their beliefs and attitudes toward new learning (Spillane, Reiser, & Reimer, 2002). This notion of teachers' attitudes toward reforms emphasizes the importance of understanding how teachers actually implement reform as intended (Henstrand, 2006, Spillane et al., 2006). Therefore, teacher buy-in is particularly critical since it is closely linked with teachers' attitudes and motivation toward instructional changes. One of the critical elements of reforming schools is related to changes in teacher beliefs "in whether or not instructional practices will be changed and how they will be implemented and sustained" (Johnson, 2006, p. 159). While research has increasingly paid attention to meaningful and effective participation in implementation processes, there is little empirical evidence about the relationship between teacher buy-in and implementation process in professional development.

Teacher buy-in could be directly linked to changes in classroom instruction and effects of professional development. However, it has been widely suggested that some teachers tend to have negative perceptions and motivations toward new educational reform policies (Leithwood, Steinbach, & Jantzi, 2002). Research suggests that more experienced teachers could be more reluctant than less experienced teachers to change their teaching styles (Huberman, 1989). For instance, Borman and his colleagues (2008) found that an inquiry-based science professional development program has a better impact on students taught by less experienced teachers than their counterparts taught by more experienced teachers. This difference in student outcomes of professional development implies that teachers may have different levels of receptivity toward implementing principles of professional development according to their teaching experience.

CHAPTER IV

VARIABLES AND MEASURES

IV. 1. Variables and Measure

This study targets the Los Angeles Unified School District (LAUSD), the nation's second-largest school district. Using stratified random sampling, eighty schools were selected from a total of 191 schools nominated by the 8 LAUSD local districts. The intervention was randomly assigned to 80 schools, including 40 treatment and 40 control schools. In the summer of 2006, 40 treatment schools were encouraged to send two teachers, including one science lead teacher (SLT) and another teacher from the same grade level, to a 5-day immersion professional development program that covered Life Science. The current study focuses on the first year of implementation with 4th grade level teachers.

78 out of 80 schools administered the 2006-07 life science periodic assessments in the original data of all study schools. Classroom observation data are available for 196 of 274 teachers (72%) in study schools for 2006-07. An analytical sample of SLTs includes 35 schools (17 in the treatment; 18 in the control schools), 38 teachers (19 in the treatment; 19 in the control schools), and 979 students (495 in the treatment; 484 in the control schools). The analytical sample of all teachers includes 48 schools (22 in the treatment; 26 in the control schools), 80 teachers (42 in the treatment; 37 in the control schools), and 4530 students (2257 in the treatment; 2273 in the control schools). As baseline statistics show in Table 2, there are no statistically significant differences between control and treatment schools.

[Table 2 here]

For the purpose of this study, the sample was selected from district administration data, classroom observation data, and teacher survey data. The data includes 4th grade students' demographic characteristics, science achievement scores, teacher observations in the classroom, etc., as presented in Table 2. Analytic data consist of two samples: 1) only those students taught by SLTs, and 2) all 4th grade level students and teachers, as shown in Table 3 and Table 4, respectively.

[Table 3 and Table 4 here]

1) Implementation Variations

To examine implementation gaps and the different impact of inquiry-based science professional development on inquiry-based instruction, three different measures were used. First, "cumulative use" describes how frequently teachers implemented each and any features in their instruction. For example, if a teacher adopted feature 1 three times and feature 2 twice and none of the others, feature 1 will be counted as 3; feature 2 as 2; features 3, 4, and 5 as 0; and "any feature" as 5. Second, "binary use" is defined as whether each and any features were implemented in instruction or not. For instance, if a teacher adopted feature 1 three times and feature 2 twice and none of the others, then feature 1 will be indicated as 1; feature 2 as 1; features 3, 4, and 5 as 0; and "any feature" as 1. Finally, "completion of cycle" is defined as the extent to which the complete cycle of five features is implemented. For example, if a teacher adopted feature 1 three times and feature 2 twice and none of the others, then the completion of cycle will be measured as 2. If a teacher adopts all five features, it will be measured as 5.

2) Variables

The independent variable is whether or not schools participated in the science immersion units (1=treatment, 0=control). This treatment variable was randomly assigned to the school level. The dependent variable of interest is 4th grade students' life science scores, which was assessed using the periodic assessment in science in the 2006-07 school year. Since the treatment was assigned to the school level and the outcome variable is at the student level, the analysis must consider multilevel structure.

The mediating variable includes cumulative use, binary use, and completion of cycle. Based on the classroom observation data, "any features of inquiry" variable is composed of each of the 5 individual features of inquiry variables as a single "any features of inquiry" variable, from the class observation data. Since those 5 features of inquiry are considered as a cycle, "any features of inquiry" is a reliable measure for the reformed teaching practices in the classroom. To obtain classroom observation data, trained observers were sent to science classes taught by SLTs and 4th grade teachers in both control and treatment schools. Each classroom observer for each of the 8 local districts recorded narrative notes in 5-minute segments. Narrative notes were coded into counts in a segment applied not only to the essential feature of inquiry, but also to its variations. To adjust the distribution of the count variable, I transformed this count variable into a logged variable and aggregated the teacher level into the school level.

The moderating variables are continuous variables that measure characteristics of the school-level professional learning community, principal support, and collective teacher buy-in using data from teacher surveys, as shown in Table 5.

[Table 5 Here]

Since individual teachers are not identified, teacher-level responses were aggregated to the school level. Based on previous literature, this study uses 6 categories with 9 total items to measure the various dimensions of the professional learning community, which are constructed from the survey questionnaires. Cronbach's alpha, which measures reliability, is 0.835 for the 12 survey items. The measure for principal support is constructed from 6 items ($\alpha = 0.891$), and a variable for the collective teacher buy-in is created from three items ($\alpha = 0.802$). The covariate variables include student pre-test scores for the science periodic assessments in the 2005-06 class year, class size, and years of teaching experience, which are constructed at the school level. A systematic focus on pretreatment covariates is significant in this study.

CHAPTER V

METHODOLOGY

V. 1. Analytical Strategies

In light of the research questions, the analytical strategies are threefold. First, this study decomposes the effect of an inquiry-based science professional development into the direct and indirect effects using causal mediation effect analysis. Second, although teachers intend to implement key principles of effective professional development, specific implementation gaps undermined and contradicted these efforts. Therefore, this study defines the implementation gap between what this professional development intended to change and what actually has been accomplished. Given the different measures of implementation, this study examines which schools and teacher characteristics contribute more to the changed instruction. The third analysis examines why the professional development yields negative student outcomes in the treatment schools. Specifically, to address what school organizational capacity can contribute to the effect of inquiry-based instruction, this study examines the interaction effect of inquiry-based instruction and school organizational capacity within the treatment schools.

V. 2. Causal Mediating Effect Analysis

> "Scientifically, mediation tells us 'how nature works' and, practically, it enables us to predict behavior under a rich variety of conditions and interventions" (Pearl, 2012, p. 1).

This study attempts to identify causal mechanisms by investigating the mediating role of teacher practice in the causal path between professional development and student outcomes. Conceptually, uncovering causal mediating mechanisms is helpful in better understanding how a treatment can be promoted by linking with a mediator (Kraemer, Wilson, Fairburn, & Agras, 2002). Methodologically, an intent-to-treatment effect model reveals the total causal effect, yet it may not be sufficient to determine the causal mechanisms (the "how"). Therefore, causal mediation analysis is becoming vital for exposing the black box in randomized experimental studies.

Traditionally, the linear structural equation model (LSEM) has been utilized in social science for mediation analysis (Imai et al., 2010; VanderWeele, 2009). However, this framework has the following limitations: 1) LSEM only allows for specific statistical models under the particular assumptions, and 2) LSEM does not suggest non-linear or nonparametric models (Imai et al., 2010; VanderWeele, 2009).

To overcome these limitations, this study adopts the average causal mediating effect model (ACME) recently developed by Imai et al. (2010). ACME proposes three statistical approaches. First, the causal mediation framework is formulated based on the counterfactual framework of causal inference, and it relies on weaker assumptions than LSEM.

Second, Imai et al. (2010) identified nonparametric models under the sequential ignorability assumption, which is critical to ACME. The first of the two sequential ignorability assumptions is that a treatment is independent of potential outcomes and mediators. In other words, the treatment assignment is independent of the given observed pretreatment confounders (X_i). Since one of the strongest design factors of this study is that the treatment was randomly assigned to the treatment and control schools, the first assumption is fulfilled. The second assumption is that the observed mediating variable is independent of all potential outcomes given the treatment and pretreatment covariates. However, since the value of the mediator variable is not randomly assigned, I control for covariates (i.e. student pretest score, class size, and teaching experience) that potentially confound the relationship between the outcome and the mediator. It is reasonable to consider pretest score, class size, and teaching experience as being predictive of both the reformed teaching practice and the student outcomes. Nonetheless, it is not fully guaranteed that these covariates are the sole confounding variables in this study. Hence, sensitivity analysis plays a critical role in testing the robustness of findings in order to address possible unmeasured confounders (Imai et al., 2010).

Third, this framework allows for linear and nonlinear relationships, parametric and nonparametric models, continuous and discrete mediators, and various types of outcome variables. In particular, this nonparametric identification requires no distributional or functional form assumptions.

The hypothesis is that participation in the inquiry-based science professional development improves student outcomes by facilitating teachers' inquiry-based science practice. The professional development was randomly assigned at the school level (j), and the outcome is each individual student level (i). M_j denotes the mediator, indicating the number of *any features of inquiry* that were observed in school j, which was measured after professional development was implemented but before measuring the outcome variables. Let T_j denote the treatment variable where $T_j = 1$ indicates that school j is assigned to the professional development (treatment), and $T_j = 0$ indicates that school j is assigned to the control group. Y_{ij} represents the outcome variable, and X_j implies the vector of observed pre-treatment covariates.

To define the causal mediation effect, the potential outcome framework is one of the key concepts. Both Y_{1ij} and Y_{0ij} can be defined as potential outcomes. If a school is assigned to the treatment condition, only Y_{1ij} is observable. Likewise, if a school is assigned to the control condition, only Y_{0ij} can be observable. In this study, the causal effect of the professional development on student outcomes can be described as the difference between two potential outcomes: one in which student outcomes would have occurred if the school had participated in the treatment and the other in which the outcomes would have occurred if the school had not participated in the treatment. Likewise, there are two potential values of a mediator, M_j (1) and M_j (0), and likewise, only one of the values can be observed.

The causal effect of treatment on student *i*'s outcome is $Y_{1ij} – Y_{0ij}$, and the average causal effect of treatment is $E[Y_{1ij} – Y_{0ij}]$. It should be noted that we assume there is no interference between treatment and control schools. Causal mediation effects for each unit *i* and *j* can be defined as:

$$\delta_{ij}(t) \approx Y_{ij}(t, M_j(1)) - Y_{ij}(t, M_j(0)), \qquad (1)$$

For each treatment and control condition, the causal mediation effect indicates the indirect effect of the professional development on student outcomes through the mediating variable of teaching practice. Equation (1) leads to the counterfactual question: What would happen to the student science outcome if a school were to change its teaching practices from the value under the control condition, $M_j(0)$, to the value under the treatment condition, $M_j(1)$, while holding the treatment status at t ? Thus, $\delta_{ij}(1)$ indicates the difference between the two potential science outcomes for students in the treatment schools. $Y_{ij}(1, M_j(1))$ indicates student *i*'s outcomes if school *j* participated in the professional development. $Y_{ij}(1, M_j(0))$ equals student *i*'s outcomes under the professional development participation, assuming mediator values as one would have observed if school *j* had been under the control condition. Likewise, $\delta_{ij}(0)$ represents student *i*'s outcomes stemming from the change in the mediator caused by participation in the professional development, while holding constant the direct effect of professional development participation. Likewise, the direct effect can be defined as follows:

$$\delta_{ij}(\mathrm{t}) \approx Y_{ij}(1, M_j(t)) - Y_{ij}(0, M_j(t)), \qquad (2)$$

$\delta_{ij}(1)$ indicates the direct effect of the professional development on student outcomes while holding the mediator constant, which would have happened under the treatment condition. The total effect is the average of the causal mediation and direct effect, as below:

$$T_i \approx Y_{ij}(1, M_j(1)) - Y_{ij}(0, M_j(0)), \qquad (3)$$

Finally, the average causal mediation effect can be indicated as following (4). For $t_{ij} = 0, 1$, this generalizes the average causal mediation effect on the outcomes of the student population.

$$\delta_{ij}(t) \approx E[Y_j(t, M_j(1)) - Y_j(t, M_j(0))], \quad (4)$$

Further, given the potential outcome framework, we can allow the causal mediation effect to vary under both condition and treatment. In particular, this study focuses on heterogeneous mediating effects of teaching practices between professional development and student outcomes. Therefore, addressing interaction terms is critical. The potential outcome framework is an important concept for understanding what interaction terms operate in this framework.

To conduct the analysis, I used *Mediation: R package for Causal Mediation Analysis,* developed by Imai, Keele, Tingley, and Yamamoto (2010). While sensitivity analysis is one of the most critical components in this analysis, the current version of the R package does not allow sensitivity analysis for multi-level regression models[2].

V. 3. Hierarchical Linear Model

First, to investigate the relationship between the implementation of the inquiry-based science professional development and the school organizational capacity, this study employs hierarchical linear modeling (HLM).

$$(\text{Inquiry use})_{ij} = \beta_{0j} + \beta_1 (T)_j + \beta_2 (SC)_j + \beta_3 C_{ij} + \beta_4 X_j + u_j + e_{ij}$$

---

[2] Fritz and MacKinnon (2007) test the ways in which mediation analysis studies can determine the adequate sample size to detect an effect. Their findings suggest that as a direct effect increases, the required sample size decreases to achieve 0.8 power in Baron and Kenny's (1986) model.

In Model 1, The binary variable T_j denotes the treatment assignment status for school *j* (T_j=0 for the control schools and T_j=1 for the treatment schools), and β_2 (SC) includes the following factors: professional learning community, principal support, and collective teacher buy-in. $\beta_3 C_{ij}$ is a vector of lesson level variables, $\beta_4 X_j$ is a vector of school-level variables including pre-test scores (the school mean percentage correct on the fourth-grade life science periodic assessment in 2005–2006); e_{ij} denotes a residual; and u_j indicates random effects for school *j*. Models 2, 3, and 4 include the variables in Model 1, as well as interaction terms between the treatment assignment and school capacity variables, in order to examine how the instruction changes according to the school organizational capacity.

As described above, this study uses three different outcomes for the use of inquiry features, including cumulative use of any inquiry features as a continuous dependent variable, binary use of any inquiry features as a binary dependent variable, and completion of cycle as an ordered categorical dependent variable. Therefore, I use a multilevel logistic regression analysis for the binary use of inquiry features. Given that the completion of cycle is an ordered categorical variable, multilevel ordered logistic regression was utilized.

Second, to determine which school and teacher characteristics deteriorate or improve student learning within the treatment group schools, several multilevel regression models were utilized.

$$(\text{Student outcome})_{ij} = \beta_{0j} + \beta_1 (SC)_j + \beta_2 C_j + \beta_3 X_j + \beta_4 Z_{ij} + u_j + e_{ij}$$

β_{0j} is an intercept, and β_{1} $(SC)_{j}$ indicates professional learning community, principal support, and collective teacher buy-in. β_{2} C_{j} is a vector of lesson level variables; β_{3} X_{j} is a vector of school-level variables including pre-test scores (the school mean percentage correct on the fourth-grade life science periodic assessment in 2005–2006); β_{4} Z_{ij} denotes student-level variables, including whether students are assigned to SLTs and novice teacher status; e_{ij} denotes a residual; and u_{j} indicates random effects for school j.

CHAPTER VI

RESULTS

VI. 1. Descriptive Statistics

1) Demographic Characteristics

As baseline statistics show in Table 2, there are no statistically significant differences between the control and treatment schools. Specifically, 76.3% and 71.2% of students qualify for the free or reduced lunch program in the treatment and control schools respectively. 33.2% and 32.3% of the students are English language learners in the treatment and control schools respectively. In the control schools, 12.4% of the students are White, and 68.2% of the students are Hispanic. In the treatment schools, 13.5% of the students are White, and 69.8% of the students are Hispanic.

[Insert Table 6 here]

Table 6 shows descriptive statistics of the main variables for teachers. In both groups, SLTs and all 4th grade teachers, there are more female teachers than male teachers. While 82% and 83% of teachers in public schools were non-Hispanic White in the 2011–12 school year and 2007–08 school year respectively (NCES 2013; 2009), this sample includes only 38.26% of all 4th grade teachers and 35.94% of SLTs are White in the control group of this study. Interestingly, the majority of teachers are Hispanic in this sample; 41.74% of all 4th grade teachers and 45.31% of SLTs are Hispanic in the control group schools. For all 4th grade teachers, 27.40% of the control school teachers and

26.60% of the treatment school teachers have above a master's degree, while amongst SLTs 21.5% of the control school teachers and 26.80% of the treatment school teachers have above a master's degree. For all 4th grade teachers, the control school teachers have 11.269 years of teaching experience and the treatment school teachers have 10.589 years of teaching experience, while the control school teachers have 10.187 years of teaching experience and the treatment school teachers have 10.863 years of teaching experience for the SLTs.

To measure teachers' professional learning community, teacher buy-in, and principal support perceived by teachers, a teacher survey administered in 2006 fall was used. The descriptive statistics show that the group of all 4th grade teachers in the treatment schools (Mean=16.807) reported a higher level of teacher professional learning community than teachers in the control group schools (Mean=16.247); however the difference is not statistically significant. SLTs in the treatment schools (Mean=16.433) reported a higher level of teacher professional learning community than SLTs in the control group schools (Mean=16.796); however the difference is not statistically significant. For both categories, all 4th grade teachers and SLTs, the teachers in the treatment schools reported a higher level of buy-in than teachers in the control schools, and the differences are statistically significant. As for teacher-reported principal support, both teacher groups in the treatment schools reported a higher rate than teachers in the control schools, yet the difference is statistically significant only for the sample including all 4th grade teachers, not for the SLTs.

2) Implementation Gaps

Tables 7, 8, and 9 show descriptive statistics for use of inquiry features for different measures, such as cumulative use, binary use, and completion of cycle.

[Tables 7, 8 and 9 Here]

Cumulative Use

The cumulative use measure indicates how many of each and any features were present in classrooms. As Table 7 shows, some inquiry features were implemented more frequently than others. Based on observations by trained researchers who counted inquiry features in the classrooms, teachers in the treatment schools more frequently used inquiry-based science features in their instruction. For both 4^{th} grade level instructors and SLTs, the difference in using any inquiry features between the treatment and control schools is statistically significant (*t*-test=4.482*** for all 4^{th} grade level instructors; *t*-test=4.209*** for SLTs). Specifically, the findings show that teachers in both treatment and control schools rarely use features 4 and 5, while features 1, 2, and 3 are more present in their practices. However, there are no different patterns in the use of features 4 and 5 between the control and treatment group schools. The differences between the treatment and control schools in using features 4 and 5 are not statistically significant. As Figure 2 also illustrates, the raw count of any feature use shows that frequency from 0 to 5 is more present in the control group, while frequency above 5 is more present in the treatment schools.

[Figure 2 Here]

Taken together, the results of cumulative use indicate that the immersion promotes teachers to used inquiry-based instruction more frequently; however, some aspects of inquiry-based instruction were more present than others in their practices.

Binary Use

The binary use measure is defined as whether each and any features were implemented or not. As shown in Table 8, the findings suggest that the immersion unit changes teachers' inquiry-based instruction only partially, which corroborates Grigg et al. (2013). The findings also show that 87% of the instruction in the treatment schools included at least one of the inquiry features, as compared to 67% of the instruction in the control schools, when all 4th grade teachers are considered. However, more in-depth examination reveals that the treatment did not impact the implementation of each feature. For instance, only 12% of the treatment instruction implemented feature 5, as compared to 11% for all 4th grade level teachers. While there are significant differences in using any features, feature 1, 2, and 3 between the control and treatment group schools for all 4th grade teachers, only "any features" and feature 1 are statistically significant for SLTs.

[Figure 3 Here]

Completion of Cycle

This immersion unit is designed to be cyclical; therefore, the evaluation should examine whether entire features were fully and cyclically implemented. This type of implementation measure, completion of cycle, considers how thoroughly teachers implemented the inquiry-based inquiry components, as intended.

As illustrated in Table 9, the results show that no treatment schools completed all five inquiry-based instructions, as the inquiry-based science professional development intended. For instance, the treatment schools have implemented more diverse inquiry features than schools in the control group. For both teacher groups, all 4th grade level teachers and SLTs, one, two, or three different features were more frequently implemented than four or five different features. However, for the three, four, and five different uses of inquiry features, there is no statistically significant difference between the treatment and control schools.

VI. 2. Analytical Results

1) How does inquiry-based instruction mediate the effect of professional development on student outcomes?

The results suggest a negative total effect of professional development on student outcomes for both SLT and all 4th grade level teachers, as prior analyses using the same data had also shown. The main purpose of this analysis is to decompose the total effect into direct and indirect effects, as shown in Tables 10, 11, and 12. The findings show that the direct effect has a negative impact on student outcomes both under the control and treatment condition and both are statistically significant. However, the average mediating effect of teacher practices on student outcomes is not statistically significant. These direct and indirect effects indicate the effect of this professional development is not mediated by inquiry-based instruction. These patterns show throughout the three different measures of implementation.

[Tables 10, 11, and 12 Here]

The main purpose of the causal mediation effect framework is to identify causal paths of the treatment. In addition to the mediation effect, the direct effect can provide a more comprehensive understanding of how the causal effect of inquiry-based instruction (mediator) on student outcomes (dependent variable) changes as function of the professional development (treatment). The effect of an exposure on an outcome when the intermediate is fixed at a specified level E[Y(1,M)-Y(0,M)] shows whether the treatment moderates the causal effect of the mediator on outcomes (Baron & Kenny, 1986; Imai, Keele, & Yamamoto, 2010).

Given that the total effect is shown to be predominantly due to the direct effect, this result may indicate that the impact of the professional development was primarily through other indicators. Education programs, including professional development, are multifaceted, and this direct effect reveals possible mechanisms that explain why the professional development did not work. These direct, indirect, and total effects indicate that the potential mediator of the impacts of a professional development, inquiry-based instruction, play little role in influencing student outcomes.

Which schools do benefit from professional development for inquiry-based instruction?

The findings regarding the relationship between the instructional changes and school organizational capacity, including professional learning community, principal support, and collective teacher buy-in, are shown in Tables 13, 14, and 15. The results also show how three different measures of implementation of the inquiry features in the classrooms interact with school organizational capacity. This study hypothesizes that school organizational capacity facilitates effective implementation of the treatment in schools and influences the effect of the professional development on inquiry-based teaching.

[Tables 13, 14, and 15]

As shown in Table 13, the treatment increases cumulative use of inquiry features in classrooms in model 1. Collective teacher buy-in has a positive relationship with cumulative use of inquiry features (Coefficient=1.388*). In model 2, schools with higher principal support demonstrate more cumulative use of inquiry features from the treatment (Coefficient=0.831*). Figure 4 also shows how the relationship between principal support and cumulative use of any inquiry features varies between the treatment and control group schools.

[Figure 4 Here]

For the binary use of inquiry feature in Table 14, Model 1, the findings from multilevel logistic analyses show that the treatment has a positive impact on the binary (Coefficient=1.169*). As Table 15 illustrates, the findings from the multilevel ordered logistic regression show that the treatment increases the inquiry features completion of cycle (Coefficient=0.877*), and collective teacher buy-in is positively associated with the inquiry features completion of cycle (Coefficient=0.473*) in Model 1. The interaction terms concerning the relationship between the assignment of treatment and principal support, professional learning community, and teacher buy-in, are not statistically significant in Models 2, 3, and 4.

[Figure 5 Here]

What school organizational capacity contributes to improving student outcomes in the treatment schools?

Given the negative total and direct effects of the professional development, this study attempts to examine which teacher and school characteristics may deteriorate or improve student outcomes in a treatment school. As shown in Table 16, Model 1, novice teachers in the treatment group have a negative relationship with student science learning outcomes. In addition, students taught by SLTs tend to have higher scores on the life science test; in Model 2, SLTs tend to contribute to students' science learning. However, the interaction effect between the cumulative use of any features and science teachers suggests that SLTs tend to have a lesser impact on students' science scores when they are more likely to use any inquiry features more frequently.

[Tables 16, 17, and 18]

With respect to binary use of any features in schools in Table 17, Model 1 suggests that students in the treatment schools that have greater professional learning community tend to have better science scores. Students taught by novice teachers tend to earn lower scores on their life science test than their counterpart students taught by more experienced teachers. In Model 2, with the interaction term between the binary use of inquiry features and SLTs, the results show that SLTs tend to produce higher scores on 4th grade students' life science tests than non-SLTs do. The finding also shows that the treatment school SLTs tend to contribute less to their students' life science test scores as they are more likely to use inquiry-based instruction. The professional learning community in the school has a positive relationship with students' outcomes (Coefficient=2.908*). In Model 3, the interaction effect between binary use of any features and novice teachers shows that students taught by the treatment school novice teachers who used inquiry-based practices benefit more on their life science test scores.

[Figures 5 and 6 Here]

Using completion of cycle for the measure of inquiry-based instruction, Table 18 shows that additional use of different inquiry features is not significantly related to students' science scores in Model 1. Within the treatment, the use of inquiry-based instruction is not significantly related to student outcomes. The higher a school's level of professional learning community, the higher the scores students earn (Coefficient=3.006*). Students taught by SLTs tend to earn higher scores than students taught by non-SLTs (Coeffcient=4.193***). In Model 3, while novice teachers tend to produce lower scores than more experienced teachers (Coefficient=-39.225***), the interaction term suggests that novice teachers tend to improve student performance more when they complete more of the cycle of inquiry features (17.553***). This result further underscores the importance of completion of the cycle of inquiry features.

CHAPTER VII

DISCUSSION

This study contributes to better understand the key mediation mechanisms involved in the relationship among teachers' professional development, reformed practices, and student outcomes. The findings help demystify the black box of how professional development impacts student outcomes via teachers' classroom practices. This research seeks to empirically test theoretical models of the impact that professional development has on both teaching and student outcomes. As new reformed practices are implemented, teachers need professional development while school leaders must manage proper school environments. This study demonstrates how to most effectively implement professional development and promote changes in instructional practices and student learning. This study will be of particular interest to educators and researchers attempting to change instructional practices in urban communities.

Educational scholars have pointed out that important questions remain unanswered regarding the mechanisms explaining how instructional changes occur and how professional learning can be linked to student learning (Knapp, 2003). Researchers have increasingly paid attention to identifying such causal mechanisms instead of focusing on whether or not programs work. Nonetheless, empirical evidence still remains scarce. First, there is little evidence about the causal mechanisms that impact the ways implementation of professional development leads to changes in both teaching and learning. Second, little is known about why some professional development programs fail to improve either teacher or student learning. For instance, this particular professional development program yields more changes in science instruction in the treatment schools compared to the control schools (Grigg et al., 2013), but the intervention has failed to improve students' test scores. This mixed finding provokes important questions, such as what kind of changes in instruction the professional development yields and why those instructional changes do not lead to improved student test scores. Third, we have limited knowledge about key links among reform initiatives, professional development, instructional change, and student learning, in spite of the important role of professional development as a policy pathway. For instance, studies often overlook to examine the ways in which teacher professional development influences teacher practices, which in turn can improve student achievement. Lastly, we lack comprehensive understanding about the implementation process and its consequences for teaching and learning. Research often overlooks how teachers adopt what they were taught from professional development, to what extent implementation gaps correlate with student learning, and what school and teacher characteristics can be associated with these gaps and changes in practices.

To address these critical gaps in the relevant literature, this study raises the following questions: 1) How do reformed classroom practices mediate the relationship between professional development and student outcomes? 2) What school and teacher characteristics are associated with inquiry-based inquiry instruction? 3) How do different levels of school factors, including professional community, principal support, and teacher buy-in, moderate the relationship between professional development and inquiry-based practices? 4) Within the treatment group schools involved with the reformed instruction, which schools and teachers contribute to student outcomes? The remainder of the section discusses the findings.

How does professional development influence instruction and student outcomes?

Given the mixed findings in prior research (Borman et al., 2008; Grigg et al., 2013), it is meaningful to decompose the total effect of the professional development into direct and indirect effects. Knowing not only whether policy interventions work, but also why and how these programs work, can enrich understanding of the intervention, which can in turn help policymakers prescribe better policy approaches or revise existing interventions (Keele, Tingley, & Yamamoto, 2015; Gamoran, 2013).

On the one hand, the result of comparing treatment and control group schools confirms the findings by Grigg et al. (2013) that the treatment increases implementation of any inquiry features although this implementation focuses on only one or two features rather than a complete cycle. On the other hand, the current study extends the measure of implementation to verify how professional development drives changes in instruction. To understand how classroom instructional changes occur through professional development, this study specified three different ways to evaluate implementation outcomes. First, the cumulative use of any features gauges how frequently inquiry-based practices occur in teaching. Second, the binary use of inquiry features assesses whether or not any inquiry features appeared in the lesson. Third, the completion of cycle measure captures how fully the implementation has been completed. These different measures of implementation can be used differently depending on the targeted outcome. For instance, if the targeted outcome was whether or not teachers completed a cyclical implementation of the five features, the treatment might have shown no effect (see Table 9; there is no statistically significant difference between the treatment and control groups for the use of three, four, or five different number of features).

Regarding the measures of implementation outcomes, this study focuses on the implementation gap and school organizational capacity that may influence the gap. First, the treatment increases instructional change in general. For instance, the treatment increases the chance that any inquiry features will appear, the average number of any inquiry features, and the number of different inquiry features. Interestingly, there is no difference between SLTs and non-SLTs in implementing the inquiry features. In particular, teachers benefit from the treatment by increasing the number of any inquiry features in their instruction when they perceive that their principals are supportive. This finding indicates that school leadership plays an important role for teachers in implementing new teaching approaches. A large body of research indicates the important role principals play in helping implement new programs in schools (Datnow & Castellano, 2001). However, there is no statistically significant interaction effect between the treatment and principal support when it comes to completing the cycle of inquiry features.

As a driver of school reforms, professional development influences teachers, and in turn impacts student outcomes through the mediating effects of instruction. The findings of this study show that inquiry-based instruction does not significantly mediate the effect of professional development on student outcomes. The results indicate that the mediation effect of inquiry-based instruction between professional development and student outcomes is positive but not significantly different in the control versus the treatment. However, the direct effect of the professional development on student outcomes is negative and significant. An ideal scenario might be one in which a positive mediation effect is larger than the negative direct effect, so that the treatment would have a positive total effect. This negative direct effect indicates that implementing the professional development can negatively influence student performance, after controlling for a mediation effect. The negative direct effect may emphasize the importance of school environment when implementing new programs, including factors such as how capably teachers and principals manage new programs. Teachers and principals may need some time to adjust to changes that may demand additional workload and possibly disturb student learning. This result further highlights the importance of preparing and monitoring school environment when implementing and evaluating new programs in schools.

Why does professional development fail to improve student outcomes?

It is less sufficient to conclude causal findings without knowing the mechanisms (Acharya, Blackwell, & Sen, 2014). In light of the negative direct effect, if the inquiry-based instruction is the same between the control and treatment, then implementing this inquiry-based science professional development can be suggested to decrease students' test scores. Therefore, this negative direct effect deserves particular attention. One of the benefits of investigation into the direct effect is when researchers attempt to rule out a mediating variable using a placebo group. For instance, when participants read articles with similar content but different format, researchers would include a control group that reads an article with entirely different content. When comparing between the control and treatment groups, researchers can examine a direct effect where "reading an article"—the mediating variable—is the same for participants. Therefore, researchers can rule out an alternative, revealing that reading an article is what changes people's attitudes rather than the content of the article (see Acharya, 2014). Another benefit of examining a direct effect is detecting whether the treatment is mediated by a particular variable, making this mediator part of a causal mechanism (Acharya, 2014). Dimitrakopoulos and Richardson (2001) argue that implementation is the process where the choice of policy instruments and resources are examined against reality, which can make it hard to achieve the goal. One reason for this negative direct effect can be related to the concept of implementation deficit; implementing new programs can be a challenge for schools, which may lead to unintended results. Both half-completed implementation and negative direct effect on student outcomes may indicate an implementation deficit or implementation dip. Dupuis and Knoepfel (2013) define implementation deficit as "the outputs of a policy failing to contribute more than symbolically to problem solving" (p. 2). In public administration literature, the concept of implementation deficit or incomplete implementation is not new, as street-level bureaucrats are understood to be overwhelmed by the competing demands of their jobs (Blackmore, 2001; Lipsky, 1980).

Similarly, a rich body of literature suggests that implementation dips are especially common during the first year of the implementation (Borman et al., 2003). Implementing a new program and attempting to transform instruction can be a challenge, which may lead to a negative impact on teaching and learning. This negative direct effect and positive but non-significant mediation effect underline the important role of actual actions from meaningful learning processes rather than implementation without comprehensive understanding about a new project.

To gauge teachers' implementation and their meaningful learning with understanding, this study uses three different ways of measuring implementation. In light of the results for the implementation of inquiry features, what and how teachers learn from professional development may not necessarily match how frequently and thoroughly they implemented what they were taught in professional development. Franke, Carpenter, Levi, and Fennema (2001) argued that when teachers learn with understanding, they should be able to apply what they learned to new topics and adjust to unfamiliar problems in their instruction. One distinguishing factor of learning with understanding is generativity, defined as "individuals' abilities to continue to add to their understanding" (Franke et al., 2001, p. 655). As learners, teachers may struggle with acquiring new approaches and knowledge. If teachers do not learn with understanding, their learning can be less comprehensive, so that the acquired skills may only be used for a particular instruction (Franke et al., 2001).

Given the striking finding that no treatment schools had fully implemented the five inquiry features, one possible reason for the non-significant mediation effect is that the professional development may not have provided meaningful learning processes for teachers. Furthermore, there is no evidence that this professional development improves teachers' content knowledge and skills, while we do have evidence that professional development changes their instructional methodologies. If the partial implementation in instruction suggests that teachers are processing changes from new approaches before the new methods become their routine in the classrooms, it may not be surprising to find that this combination weakens student learning. Therefore, this study describes the gap between what teachers learned from professional development and what they actually implemented. Although teachers in the treatment schools more frequently used inquiry features in instruction, this does not necessarily indicate their learning with comprehensive understanding about inquiry-based science instruction. Therefore, the findings highlight the importance of better focus on measures for valid implementation.

What aspects of schools and teachers contribute to the effects of professional development?

Implementation is a complex and multifaceted process of putting programs or policies into practice via mechanisms that involve a variety of conditions and actors (Dimitrakopoulos & Richardson, 2001). When increased educational resources do not improve students' learning, an important question arises: whether the resources can be better focused (Cohen et al., 1998). As Fullan (1991) noted, "staff development cannot be separated from school development" (p. 331). Furthermore, implementing transformed instruction in daily practice can be challenging and can vary according to school environment and resources that support teacher learning (Spillane, Diamond, Walker, Halverson, & Jita, 2001). Therefore, to better implement educational resources, such as professional development, it is important to identify which school organizational conditions are related to the effect of professional development. The remaining questions include: 1) what organizational factors can be promoted to improve implementation? and 2) what organizational factors and teacher characteristics are related to student outcomes when implementing inquiry features in the treatment schools?

To respond to these questions, the findings of this study suggest that teachers in the treatment schools with better principal support more frequently use inquiry features in their instruction. Datnow and Castellano (2001) address the role of leadership underpinning the relationship between teachers and principals during implementation of the Success for All program. The findings of the current study also indicate that collective buy-in is positively related to completion of the cycle of inquiry features. Not surprisingly, teachers' positive receptivity towards programs is highly correlated with how they respond to implementation of programs (Datnow & Castellano, 2000; Fullan, 1991).

With respect to the relationship between inquiry-based instruction and students' science performance in the treatment schools, the results suggest that students taught by teachers with fewer than 3 years of teaching experience obtain lower scores than students taught by more experienced teachers. Studies suggest that less experienced teachers can be less effective than more experienced teachers (Darling-Hammond, 2000; Rockoff, 2003; Rosenholtz, 1986). However, students of novice teachers benefit more from the use of inquiry-based instruction than students of more experienced teachers. This finding corroborates research suggesting that less experienced teachers are more willing and able to implement new programs in schools (Ghaith & Yaghi, 1997).

CONCLUSION

This research attempts to unpack the mechanisms of professional development by decomposing the total effect into the direct and indirect effects. Such examinations should be a constitutive feature rather than an "add-on" in education evaluation research

(Hong, 2012). In spite of educators' best intentions, programs and policy initiatives often yield unintended outcomes. Given that the intervention does positively impact the mediator but also has unexpected results on student outcomes, it is vital to understand the causal mechanisms. Therefore, this study demonstrates how professional development influences student outcomes as mediated by teachers' transformed instruction, and also why this effort has yielded unexpected outcomes.

First, to obtain comprehensive perspectives regarding the mechanisms of professional development, this study explores implementation of professional development and examines school organizational capacity. In spite of the negative direct and total effects of this inquiry-based science professional development, this study does not try to discourage schools from implementing inquiry-based science professional development for teachers, but rather to encourage a stronger focus on teachers' learning with comprehensive understanding.

Second, knowing targeted outcomes is critical to yield accurate evaluations. This study provides insight into which implementation elements need to be considered in order to understand the mechanisms of professional development. To capture implementation processes and gaps, this study identifies different measurements of implementation. For instance, the results indicate that the treatment effect on implementation of inquiry features can be differently determined according to the targeted outcomes (i.e. frequency or completion of a five inquiry feature cycle).

Third, from both design and evaluation perspectives, it might be worthwhile for professional development programs to include both teacher knowledge and teaching methodology components so that educators can learn with understanding. Further, researchers need to consider both content knowledge and teaching methodologies when evaluating professional development.

Fourth, the findings from this study suggest that organizational conditions and school capacity play a key role in improving implementation of professional development. This result underscores that implementing new instructional practices is multilayered with school resources and organizational capacity. To reduce the implementation deficit, cohesive support from school organizations and leadership is vital. Therefore, when implementing professional development, researchers need to include school organizational capacity and leadership in their evaluation components.

Beyond estimating average treatment effects, this study methodologically engages with the issues of how assigned treatments can be mediated by a critical conveyer using the average causal mediation effect framework.

APPENDIX A

TABLES

Table 1. Essential Features of Inquiry and Their Variations

Essential Feature	Variation A	Variation B	Variation C	Variation D
1. Learner engages in scientifically oriented questions	Learner poses a question	Learner selects among questions, poses new questions	Learner sharpens or clarifies question provided by teacher, materials, or other source	Learner engages in question provided by teacher, materials, or other source
2. Learner gives priority to evidence in responding to questions	Learner determines what constitutes evidence and collects it	Learner directed to collect certain data	Learner given data and asked to analyze	Learner given data and told how to analyze
3. Learner formulate explanations from evidence	Learner formulates explanation after summarizing evidence	Learner guided in process of formulating explanations from evidence	Learner given possible ways to use evidence to formulate explanation	Learner provided evidence & told how to use evidence to formulate explanation
4. Learner connects explanations to scientific knowledge	Learner independently examines other resources and forms the links to explanations	Learner directed toward areas and sources of scientific knowledge	Learner given possible connections	
5. Learner communicates and justifies explanations	Learner forms reasonable and logical argument to communicate explanations	Learner coached in development of communication	Learner provided broad guidelines to use sharpen communication	Learner given steps and procedures for communicate

Source: National Research Council (2000). Inquiry and the National Science Education Standards: A Guide for Teaching and Learning. Table 2-6. p.29.

Table 2. Comparison of Baseline Characteristics of Treatment and Control Schools

	Control	Treatment	F	P
Pretest	.585 (.104)	.530 (.125)	.041	.839
% Male	.63 (.490)	.68 (.474)	.845	.361
% Free lunch	.763 (.259)	.712 (.284)	1.667	.200
% English language learners	.332 (.196)	.323 (.195)	.070	.792
% African American	.095 (.154)	.098 (.137)	.034	.855
% White	.124 (.211)	.135 (.216)	.504	.472
% Pacific Islander	.006 (.024)	.005 (.014)	.523	.472
% Asian	.037 (.058)	.033 (.059)	.076	.784
% Hispanic	.682 (.272)	.698 (.289)	.303	.583

Note. Standard Deviations are in parentheses.

Table 3. Analytic Data for All 4th Grade Teachers

	Original			All 4th grade teachers		
	Control	Treatment	Total	Control	Treatment	Total
Schools	39	39	78	26	22	48
Teacher	39	43	82	38	42	80
Observation	94	102	196	91	99	190
Students	3211	3656	6867	2273	2257	4530

Table 4. Analytic Data for SLTs Only

	Original			SLTs		
	Control	Treatment	Total	Control	Treatment	Total
Schools	32	35	67	18	17	35
Teacher	23	22	45	19	19	38
Observation	52	51	103	43	45	88
Students	886	1053	1939	484	495	979

Table 5. Items for School Organizational Capacity

	Category	Items
Professional Learning Community (α = 0.835)	Reflective dialogue or professional inquiry	Teachers in this school are continually learning and seeking new ideas.
		In this school I am encouraged to experiment with my teaching.
	Deprivated practice	Teachers in this school regularly observe each other teaching science.
	Collaboration	There is a great deal of cooperative effort among the staff members.
		I make a conscious effort to coordinate the content of my courses with that of other teachers.
	Shared norms	Most of my colleagues share my beliefs and values about what the central mission of the school should be.
		A vision for student learning in science is shared by most staff in this school.
		Teachers in this school trust each other.
	Collective responsibility	Teachers in this school exhibit a focused commitment to student learning in science.
Principal supports (α =0.891)		The principal does a poor job of getting resources for this school. (Reverse coded)
		The principal encourages teachers to thoroughly teach the science content standards.
		It's okay in this school to discuss feelings, worries, and frustrations with the principal.
		The principal takes personal interest in the professional development of the teachers.
		The principal is an effective manager who makes the school run smoothly.
		The principal knows what kind of school he/she wants and has communicated this to the staff
Collective teacher buy-in (α =0.802)		Teachers here help one another put new ideas from professional development to use.
		Most professional development in this school enables us to build on our teaching experiences.
		In this school, professional development activities are well matched to teachers' needs.

Table 6. Descriptive Statistics

		All 4th grade teachers			SLTs		
		Control	Treatment		Control	Treatment	
		Frequency (%)		*t*-test	Frequency (%)		*t*-test
Teacher gender	Female	79 (68.70%)	73 (62.39%)	2.498*	46 (71.88%)	18 (28.13%)	1.378
	Male	36 (31.30%)	44 (37.61%)	2.498*	39 (59.09%)	27 (40.91%)	1.378
Teacher race	Asian	9 (7.83%)	9 (7.69 %)	0.552	4 (6.25%)	6 (9.09%)	0.390
	Black	12 (10.43%)	5 (4.27%)	2.128*	7 (10.94%)	5 (7.58%)	0.658
	Hispanic	48 (41.74%)	43 (36.75%)	0.029	29 (45.31%)	29 (43.94%)	0.194
	White	44 (38.26%)	56 (47.86%)	0.915	23 (35.94%)	24 (36.36%)	0.404
	Others	2 (1.74%)	4 (3.42%)	0.022	1 (1.56%)	2 (3.03%)	0.022
		Mean		*t*-test	Mean		*t*-test
Masters degree or above		0.274	0.266	0.118	0.215	0.268	0.097
Total teaching experience		11.269	10.589	0.329	10.187	10.863	0.177
Teacher buy in		3.960	4.682	4.147***	3.991	4.759	3.097*
School PLC		16.247	16.807	1.587	16.433	16.796	0.958
Principal Support responded by teachers		10.574	11.398	2.945*	10.709	11.460	1.884

Table 7. Descriptive Results of Cumulative Use

Status		Obs.	Any Feature Mean	Any Feature Std.	*t*-test	Feature 1 Mean	Feature 1 Std.	*t*-test	Feature 2 Mean	Feature 2 Std.	*t*-test	Feature 3 Mean	Feature 3 Std.	*t*-test	Feature 4 Mean	Feature 4 Std.	*t*-test	Feature 5 Mean	Feature 5 Std.
ll 4th	Con	94	4.457	4.500	4.482***	.872	2.017	2.723**	2.670	3.038	2.305*	.648	1.493	2.835**	.074	.445	0.588	.191	.643
	Treat	102	7.343	4.504		1.872	2.987		3.725	3.345		1.431	2.258		.117	.568		.196	.718
SLT	Con	52	4.288	3.821	4.209***	.903	1.806	2.751**	2.769	2.811	2.112*	.346	1.100	1.986*	.115	.582	0.311	.153	.500
	Treat	51	7.784	4.579		2.215	2.914		4.156	3.791		1.019	2.177		.156	.758		.235	.907

Table 8. Descriptive Results of Binary Use

Status		Obs.	Any Feature Mean	Any Feature Std.	*t*-test	Feature 1 Mean	Feature 1 Std.	*t*-test	Feature 2 Mean	Feature 2 Std.	*t*-test	Feature 3 Mean	Feature 3 Std.	*t*-test	Feature 4 Mean	Feature 4 Std.	*t*-test	Feature 5 Mean	Feature 5 Std.
ll 4th	Con	94	.670	.472	3.478***	.276	.449	2.841**	.574	.497	2.080*	.234	.425	2.400*	.042	.202	0.789	.106	.309
	Treat	102	.872	.335		.470	.501		.715	.453		.392	.490		.068	.254		.117	.323
SLT	Con	52	.711	.457	2.179*	.307	.466	2.739**	.634	.486	0.763	.173	.382	1.008	.057	.235	0.024	.096	.297
	Treat	51	.882	.325		.568	.500		.705	.460		.254	.440		.058	.237		.117	.325

Table 9. Descriptive Results of Completion of Cycle

The number of different inquiry features	All 4th					SLTs				
	Control		Treat			Control		Treat		
	Freq.	%	Freq.	%	*t*-test	Freq.	%	Freq.	%	*t*-test
0	31	32.98	13	12.75	3.285**	15	28.85	6	11.76	2.078*
1	27	28.72	24	23.53	0.784	17	32.69	12	23.53	0.905
2	25	26.6	45	44.12	2.443*	13	25	26	50.98	2.555*
3	6	6.38	14	13.73	1.822	5	9.62	5	9.8	0.098
4	4	4.26	6	5.88	0.050	2	3.85	2	3.92	0.060
5	1	1.06	0	0	1.015	0	0	0	0	-
Total	94	100	102	100		52	100	51	100	
Mean	1.234		1.764		3.370***	1.269		1.705		2.154*

Table 10. Results of Mediation effect: Cumulative Use

	All 4th grade teachers				SLT			
	Estimate	95% CI		p-value	Estimate	95% CI		p-value
Mediation Effect_0	0.009	-0.844	1.136	0.90	2.284	-2.679	7.701	0.30
Mediation Effect_1	0.221	-1.885	2.681	0.84	-1.266	-5.804	1.512	0.54
Direct Effect_0	-6.245	-11.152	-2.568	0.00	-7.895	-14.764	0.571	0.08
Direct Effect_1	-6.032	-11.061	-2.249	0.00	-11.445	-20.114	-1.286	0.04
Total Effect	-6.023	-11.179	-2.028	0.00	-9.161	-16.289	-2.202	0.02
Proportion via Mediation_0	-0.001	-0.302	0.158	0.90	-0.212	-1.547	0.451	0.32
Proportion via Mediation_1	-0.009	-1.034	0.234	0.84	0.086	-0.294	0.991	0.56
Mediation Effect (Ave.)	0.115	-1.016	1.447	0.80	0.508	-2.119	3.157	0.68
Direct Effect (Ave.)	-6.138	-11.140	-2.574	0.00	-9.670	-17.185	-2.208	0.02
Proportion via Mediation (Ave.)	-0.005	-0.570	0.130	0.80	-0.063	-0.659	0.364	0.70

Table 11. Results of Mediation effect: Binary Use

	All 4th grade teachers				SLT			
	Estimate	95% CI		p-value	Estimate	95% CI		p-value
Mediation Effect_0	-0.047	-0.966	0.985	0.88	-0.289	-3.583	2.766	0.70
Mediation Effect_1	-0.237	-2.627	1.808	0.74	0.263	-1.659	4.280	0.90
Direct Effect_0	-5.735	-11.575	-0.764	0.00	-8.911	-16.082	-1.253	0.02
Direct Effect_1	-5.925	-11.623	-1.736	0.02	-8.358	-15.666	0.179	0.06
Total Effect	-5.973	-11.507	-1.662	0.02	-8.648	-15.736	-0.573	0.04
Proportion via Mediation_0	-0.001	-0.249	0.225	0.90	0.012	-0.619	1.011	0.70
Proportion via Mediation_1	0.037	-0.444	0.676	0.72	-0.007	-0.960	0.397	0.94
Mediation Effect (Ave.)	-0.142	-1.660	1.233	0.70	-0.013	-1.347	1.937	0.94
Direct Effect (Ave.)	-5.830	-11.272	-1.550	0.00	-8.635	-15.646	-0.575	0.02
Proportion via Mediation (Ave.)	0.018	-0.332	0.445	0.68	0.002	-0.264	0.281	0.94

Table 12. Results of Mediation Effect: Completion of Cycle

	All 4th grade teachers				SLT			
	Estimate	95% CI		p-value	Estimate	95% CI		p-value
Mediation Effect_0	0.354	-1.011	2.490	0.68	0.466	-1.743	3.582	0.76
Mediation Effect_1	0.310	-1.302	2.127	0.78	-0.669	-4.578	2.101	0.62
Direct Effect_0	-6.292	-10.538	-1.048	0.00	-9.037	-16.640	-0.945	0.04
Direct Effect_1	-6.336	-10.603	-0.677	0.02	-10.174	-17.719	-2.249	0.02
Total Effect	-5.982	-10.062	-0.772	0.02	-9.707	-16.903	-1.271	0.04
Proportion via Mediation_0	-0.020	-0.580	0.241	0.70	-0.017	-0.366	0.267	0.76
Proportion via Mediation_1	-0.008	-0.849	0.306	0.80	0.049	-0.353	0.431	0.62
Mediation Effect (Ave.)	0.332	-0.553	2.499	0.66	-0.101	-2.114	1.579	0.98
Direct Effect (Ave.)	-6.314	-10.680	-1.021	0.00	-9.605	-17.001	-1.426	0.02
Proportion via Mediation (Ave.)	-0.014	-0.429	0.140	0.68	0.016	-0.196	0.203	0.98

Table 13. Inquiry Features and School Organizational Capacity: Cumulative Use

Variable	Model 1		Model 2		Model 3		Model 4	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Treatment	2.074*	0.877	-7.074	4.172	1.177	6.068	3.488	3.509
Principal support	-0.128	0.294	-0.533	0.325	-0.137	0.301	-0.103	0.301
Professional learning community	-0.286	0.317	-0.272	0.291	-0.305	0.342	-0.319	0.328
Teacher buy-in	1.388*	0.538	1.491**	0.493	1.411*	0.558	1.506*	0.609
Pretest score	0.049	0.069	0.059	0.062	0.048	0.069	0.052	0.069
Class size	-0.005	0.094	-0.010	0.092	-0.004	0.094	-0.010	0.095
% of teachers with above Master's degree	1.971	1.911	1.270	1.759	1.882	2.000	2.182	1.983
SLT	0.293	0.663	0.337	0.651	0.300	0.665	0.291	0.663
Treatment* Principal support			0.831*	0.371				
Treatment* Professional learning community					0.054	0.366		
Treatment* Teacher buy-in							-0.333	0.800
Constant	1.486	4.290	4.844	4.212	1.867	4.985	1.224	4.344
Random Effect								
School level variance	3.439	1.971	2.007	1.771	3.429	1.968	3.500	1.988
Student level residual	15.714	2.006	16.226	2.112	15.717	2.006	15.661	2.001

Note: *p<0.05 **p<0.01***p<0.001

Table 14. Inquiry Features and School Organizational Capacity using Multilevel Logistic Regression: Binary Use

Variable	Model 1		Model 2		Model 3		Model 4	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Treatment	1.169*	0.531	-3.424	2.808	-0.562	3.656	4.305	2.274
Principal support	-0.252	0.182	-.434*	0.209	-0.264	0.181	-0.211	0.184
Professional learning community	0.041	0.191	0.055	0.181	0.012	0.199	-0.034	0.202
Teacher buy-in	0.291	0.320	0.353	0.309	0.325	0.326	0.549	0.374
Pretest score	0.054	0.041	0.061	0.040	0.053	0.041	0.059	0.041
Class size	-0.068	0.067	-0.073	0.068	-0.068	0.067	-0.082	0.068
% of teachers with above Master's degree	0.671	1.208	0.390	1.156	0.527	1.238	1.136	1.247
SLT	0.491	0.433	0.513	0.423	0.507	0.433	0.493	0.434
Treatment* Principal support			0.409	0.249				
Treatment* Professional learning community					0.106	0.222		
Treatment* Teacher buy-in							-0.739	0.515
Constant	-0.016	2.671	1.224	2.702	0.550	2.919	-0.257	2.675
Random Effect								
School level variance	0.718	0.738	0.479	0.687	0.690	0.733	0.669	0.701

Note: *p<0.05 **p<0.01***p<0.001

Table 15. Inquiry Features and School Organizational Capacity using Multilevel Ordered Logistic Regression: Completion of Cycle

Variable	Model 1		Model 2		Model 3		Model 4	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Treatment	0.877*	0.382	-1.651	1.810	-1.443	2.631	1.453	1.540
Principal support	-0.105	0.124	-0.223	0.142	-0.129	0.124	-0.096	0.128
Professional learning community	-0.091	0.143	-0.094	0.134	-.0152	0.156	-.0101	0.146
Teacher buy-in	0.473*	0.237	0.514*	0.225	0.537*	0.243	0.517	0.265
Pretest score	0.025	0.030	0.028	0.028	0.023	0.029	0.026	0.030
Class size	-0.041	0.043	-0.043	0.041	-.0038	0.042	-0.044	0.043
% of teachers with above Master's degree	1.456	0.841	1.212	0.800	1.229	0.855	1.547	0.881
SLT	0.078	0.304	0.083	0.296	0.092	0.302	0.080	0.305
Treatment* Principal support			0.230	0.162				
Treatment* Professional learning community					0.142	0.159		
Treatment* Teacher buy-in							-0.136	0.352
Cut1	-0.791	1.876	-1.861	1.906	-1.872	2.189	-0.671	1.917
Cut2	0.660	1.885	-0.429	1.914	-0.429	2.196	0.786	1.929
Cut3	2.713	1.902	1.601	1.931	1.620	2.207	2.838	1.947
Cut4	4.025*	1.929	2.896	1.959	2.932	2.226	4.150*	1.975
Cut5	6.314**	2.155	5.168*	2.184	5.220*	2.423	6.443**	2.198
Random Effect								
School level variance	0.482	0.437	0.309	0.406	0.420	0.421	0.502	0.447

Note: *p<0.05 **p<0.01***p<0.001

Table 16. Student Outcomes within the Treatment Group: Cumulative Use

Variable	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Novice	-8.885***	2.106	-8.413***	2.096	-5.486	9.517	-8.898***	2.106	-8.875***	2.106	-8.883***	2.106
Pretest score	0.796*	0.404	0.811*	0.386	0.789	0.404	0.766	0.403	0.810*	0.403	0.831*	0.405
Cumulative use	-0.600	0.942	0.020	0.904	-0.588	0.943	-4.194	7.163	3.503	8.542	-2.912	4.269
Principal support	-0.596	1.790	-0.322	1.699	-0.609	1.789	-3.054	5.161	-0.517	1.783	-0.609	1.763
Professional learning community	2.701	1.542	2.511	1.465	2.703	1.541	2.698	1.524	4.608	4.237	2.658	0.521
Teacher buy-in	-1.156	2.918	-1.441	2.770	-1.131	2.918	-1.163	2.884	-1.078	2.899	-5.486	8.335
SLT	4.196***	1.158	21.726***	3.958	4.187***	1.158	4.189***	1.158	4.179***	1.158		
Cumulative use *SLT			-2.215***	0.477								
Cumulative use *Novice					-0.501	1.369						
Cumulative use * Principal support							0.295	0.583				
Cumulative use * Professional learning community									-0.245	0.509		
Cumulative use * Teacher buy-in											0.517	0.929
Constant	-20.694	23.795	-25.124	22.600	-20.437	23.798	10.727	66.166	-54.700	74.364	-2.660	39.855
Random Effect												
School level variance	88.216	31.541	78.607	27.678	88.157	31.531	85.891	31.040	86.720	31.117	85.336	30.903
Student level residual	414.406	13.996	409.949	13.841	414.377	13.995	414.471	14.000	414.431	13.998	414.489	14.001

Note: *p<0.05 **p<0.01***p<0.001

Table 17. Student Outcomes within the Treatment Group: Binary Use

Variable	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E
Novice	-8.859***	2.106	-9.034***	2.085	-61.903**	18.017	-8.852***	2.106	-8.860***	2.106	-8.863***	2.10
Pretest score	0.664	0.388	0.704	0.380	0.708	0.391	0.683	0.388	0.709	0.386	0.645	0.39
Binary use	11.329	19.692	26.342	19.409	8.533	19.911	86.870	126.053	170.522	184.759	33.970	105.7
Principal support	-1.178	1.688	-0.886	1.650	-1.124	1.705	4.587	9.658	-1.181	1.657	-1.175	1.69
Professional learning community	3.033*	1.457	2.908*	1.425	3.040*	1.471	3.058*	1.452	11.865	10.311	3.132*	1.52
Teacher buy-in	-0.452	2.999	-0.892	2.933	-0.573	3.029	-0.266	3.003	-0.176	2.958	3.852	19.9
SLT	4.189***	1.158	55.393***	8.616	4.287***	1.155	4.195***	1.158	4.169***	1.158	4.186***	1.15
Binary use *SLT			-59.012***	9.841								
Binary use *Novice					61.696**	20.815						
Binary use * Principal support							-6.452	10.642				
Binary use * Professional learning community									-9.753	11.274		
Binary use * Teacher buy-in											-4.935	22.6
Constant	-30.170	27.479	-44.733	26.969	-30.530	27.761	-100.012	118.404	-178.212	173.245	-50.562	97.3
Random Effect												
School level variance	85.491	31.276	81.488	28.876	87.452	31.776	84.796	30.904	82.078	30.077	85.938	31.4
Student level residual	414.572	14.006	406.482	13.726	412.408	13.932	414.523	14.003	414.587	14.006	414.537	14.0

Note: *p<0.05 **p<0.01***p<0.001

Table 18. Student Outcomes within the Treatment Group: Completion of Cycle

Variable	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E
Novice	-8.875***	2.106	-8.879***	2.106	-39.225***	8.531	-8.836***	2.105	-8.869***	2.105	-8.899***	2.10
Pretest score	0.737	0.393	0.731	0.395	0.768	0.392	0.848*	0.396	0.920*	0.426	0.690	0.39
Completion cycle	-1.304	4.450	-1.635	4.511	-1.94	4.452	36.982	28.774	42.346	41.515	25.822	25.1
Principal support	-1.018	1.692	-1.049	1.704	-1.060	1.691	5.433	5.076	-1.155	1.675	-0.961	1.68
Professional learning community	3.006*	1.485	3.003*	1.495	3.087*	1.485	3.118*	1.465	8.792	5.672	3.414*	1.52
Teacher buy-in	-1.154	2.977	-1.120	2.998	-1.139	2.976	-1.457	2.939	-1.893	3.024	9.168	9.85
SLT	4.193***	1.158	1.323	4.972	4.296***	1.154	4.244***	1.158	4.213***	1.158	4.209***	1.15
Completion cycle *SLT			1.587	2.670								
Completion cycle *Novice					17.553***	4.783						
Completion cycle * Principal support							-3.465	2.577				
Completion cycle * Professional learning community									-2.734	2.588		
Completion cycle * Teacher buy-in											-6.043	5.50
Constant	-20.022	24.841	-18.839	25.092	-21.625	24.834	-98.806	63.533	-118.465	96.405	-71.474	53.0
Random Effect												
School level variance	89.485	32.115	90.829	32.762	89.474	31.983	86.346	30.377	86.869	30.880	88.478	31.2
Student level residual	414.415	13.998	414.263	13.994	411.265	13.891	414.157	13.984	414.291	13.991	414.184	13.9

Note: *p<0.05 **p<0.01***p<0.001

APPENDIX B.

FIGURES

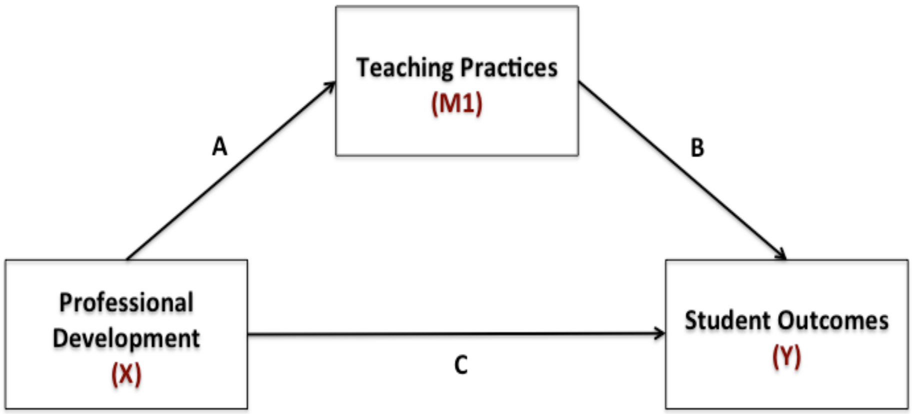

Figure 1. The Mediating Relationship between Professional Development and Student Outcomes

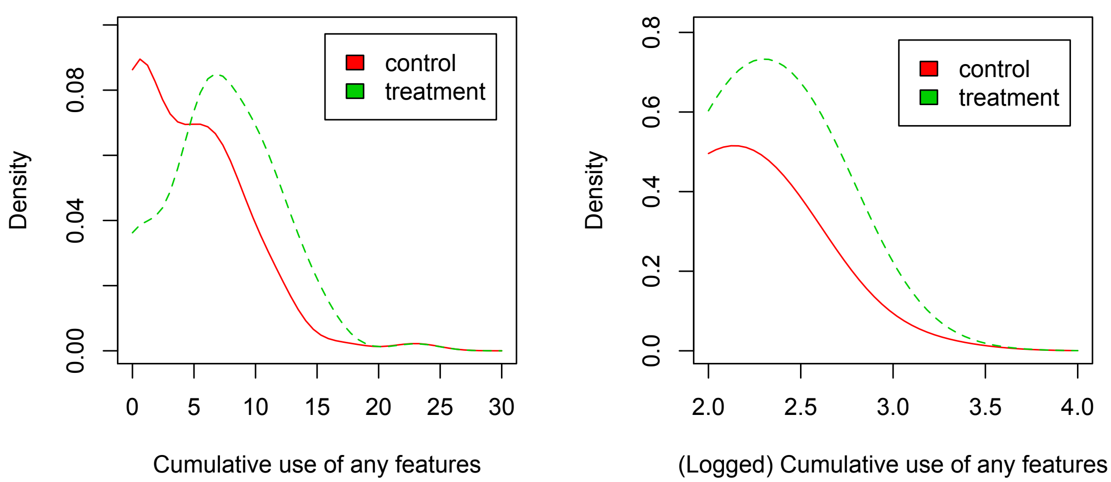

Figure 2. Description of Implementation Variation

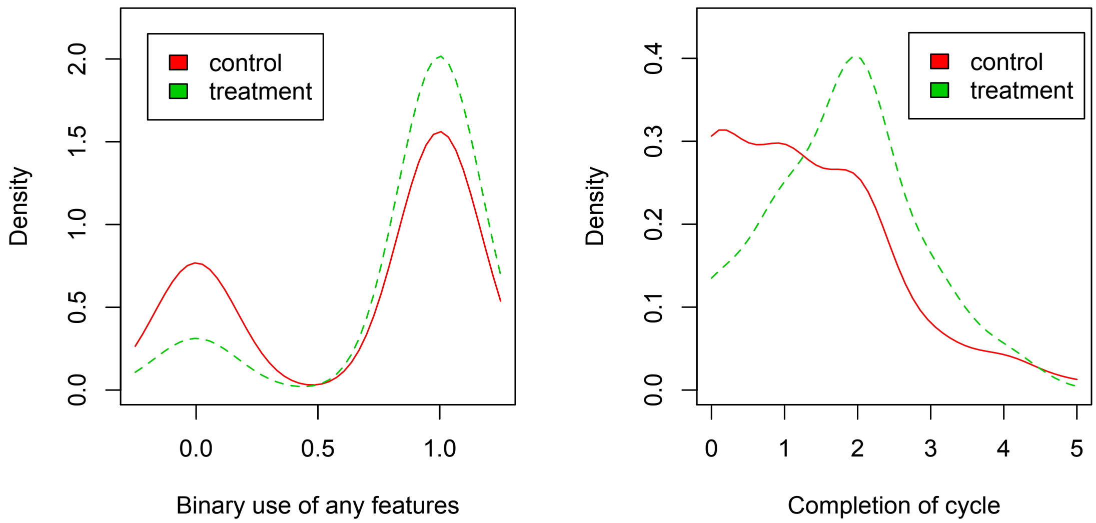

Figure 3. Description of Implementation Variation

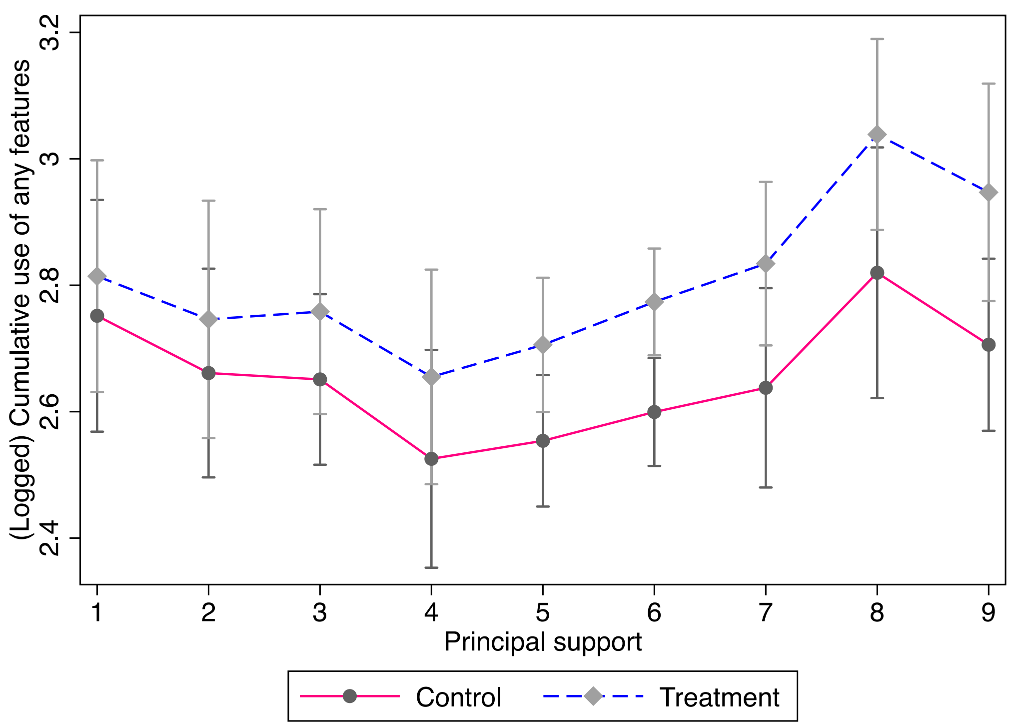

Figure 4. Interaction Effect between Principal Support and Treatment

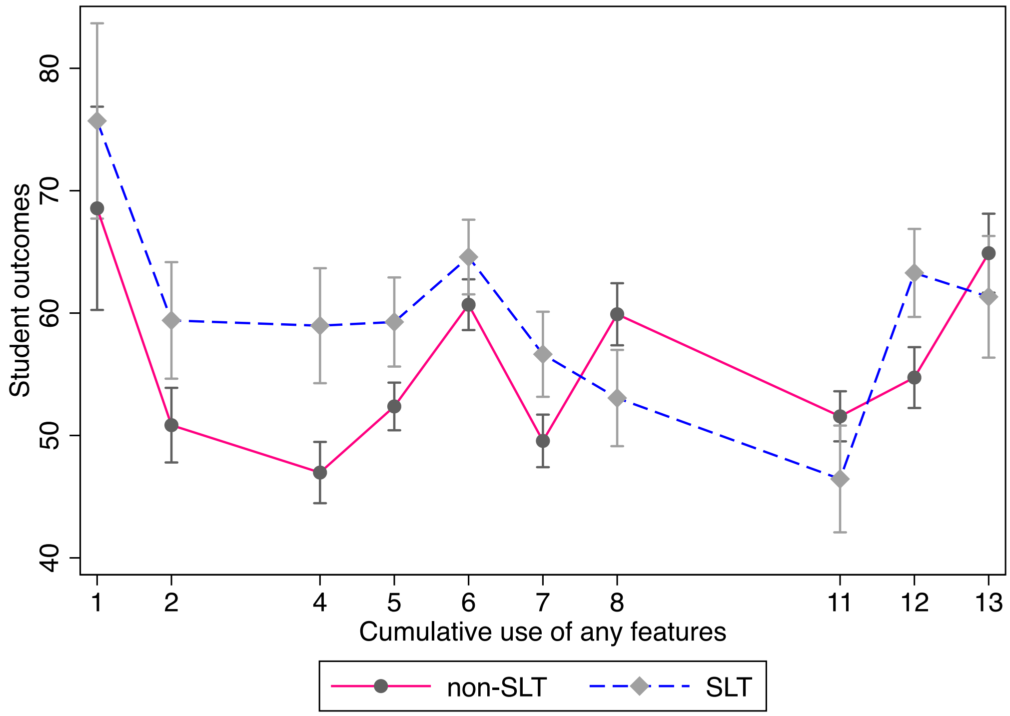

Figure 5. The Relationship between Cumulative Use and Student Outcomes

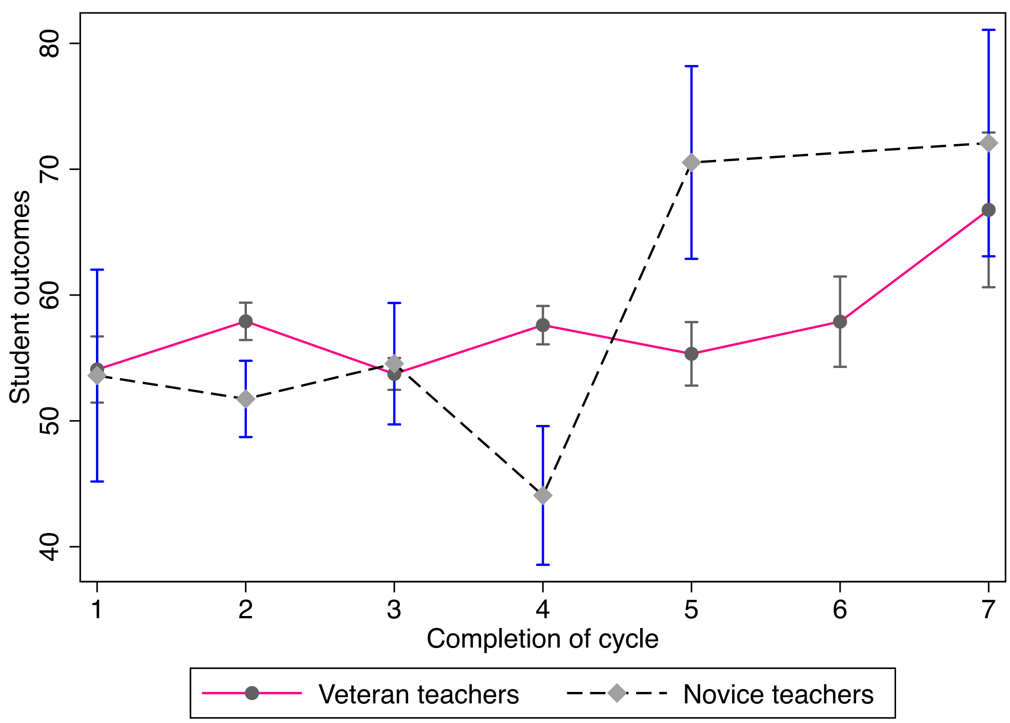

Figure 6. The Relationship between Completion Cycle and Student Outcomes

REFERENCES

Akerson, V.L., & Hanuscin, D.L. (2007). Teaching nature of science through inquiry: Results of a 3-year professional development program. *Journal of Research in Science Teaching, 44*(5), 653-680.

Anderson, R.D. (1996). Study of curriculum reform. Vol.1: Findings and conclusions. Studies of Educational Reform. (ERIC Document Reproduction Service No. ED 397535)

Avalos, B. (2011). Teacher professional development in teaching and teacher education over ten years. *Teaching and Teacher Education, 27*, 10-20.

Bascia (2010). Reducing Class Size: What Do We Know?. Retrieved from http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED420108

Ball, D.L. (1996). Teacher learning and the mathematics reforms: What we think we know and what we need to learn. *Phi Delta Kappan, 77,* 500–508.

Banilower, E., Smith, P. S., Weiss, I. R., & Pasley, J. D. (2006). The status of K- 12 science teaching in the United States: Results from a national observation survey. In D. Sunal & E. Wright (Eds.), *The impact of the state and national standards on K-12 science teachin*g (pp. 83 – 122). Greenwich, CT: Information Age Publishing.

Banilower, E. R., Heck, D. J., & Weiss, I. R. (2007). Can professional development make the vision of the standards a reality? The impact of the national science foundation's local systemic change through teacher enhancement initiative. *Journal of Research in Science Teaching*, *44*(3), 375–395. doi:10.1002/tea.20145

Borko, H. (2004). Professional Development and Teacher Learning: Mapping the Terrain. *Educational Researcher*, *33*(8), 3–15. doi:10.3102/0013189X033008003

Borman, G.D., Gamoran, A., & Bowdon, J. (2008). A randomized trial of teacher development in elementary science: First-year achievement effects. *Journal of Research on Educational Effectiveness, 1*, 237-264.

Bransford, J. D., Vye, N. J., Stevens, R., Kuhl, P., Schwartz, D., Bell, P., Meltzoff, A., Barron, B., Pea, R., Reeves, B., Roschelle, J., & Sabelli, N. (2006). Learning theories and education: Toward a decade of synergy. In P. Alexander & P. Winne (Eds.), Handbook of educational psychology (Volume 2). Mahwah, NJ: Erlbaum. Brown, J. S., Collins, A., & Duguid, P. (1989). Situated Cognition and the Culture of Learning. *Educational Researcher*, *18*(1), 32–42. doi:10.3102/0013189X018001032

Buczynski, S., & Hansen, C. B. (2010). Impact of professional development on teacher practice: Uncovering connections. *Teaching and Teacher Education, 26,* 599–607.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review*, *26*(6), 673–682. doi:10.1016/j.econedurev.2007.10.002

Coburn, C. E., & Stein M.K. (2006). Communities of Practice Theory and the Role of Teacher Professional Community in Policy Implementation. In *New Directions in Education Policy Implementation: Confronting Complexity,* ed. Meredith I. Honig. Albany: SUNY Press.

Coburn, C. E. (2001). Collective Sensemaking about Reading: How Teachers Mediate Reading Policy in Their Professional Communities. *Educational Evaluation and Policy Analysis*, *23*(2), 145–170. doi:10.3102/01623737023002145

Coleman, J. S. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology*, *94*, S95–S120. doi:10.2307/2780243

Community for Advancing Discovery Research in Education (2011). Preparing and Supporting STEM Educators, Retrieved from http://successfulstemeducation.org/sites/successfulstemeducation.org/files/Preparing%20Supporting%20STEM%20Educators_FINAL.pdf

DuFour, R. (2004). What is a ‘‘Professional Learning Community’’? [electronic version]. *Educational Leadership, 61*(8), 6.

Dubner, J., Silverstein, S. C., Carey, N., Frechtling, J., Busch-Johnsen, T., Han, J., Ordway, G., et al. (2001). Evaluating science research experience for teachers programs and their effects on student interest and academic performance: A preliminary report of an ongoing collaborative study by eight programs. *materials research society symposium proceedings* (Vol. 684, p. 73). Retrieved from http://journals.cambridge.org/production/action/cjoGetFulltext?fulltextid=821207

Fishman, B. J., Marx, R. W., Best, S., & Tal, R., T. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and Teacher Education*, *19,* 643–658.

Fleischman, H. L., Hopstock, P. J., Pelczar, M. P., & Shelley, B. E. (2010). *Highlights from PISA 2009: Performance of U.S. 15-Year-Old Students in Reading, Mathematics, and Science Literacy in an International Context. NCES 2011-004*.

National Center for Education Statistics. Available from: ED Pubs. P.O. Box 1398, Jessup, MD 20794-1398. Tel: 877-433-7827; Web site: http://nces.ed.gov/. Retrieved from http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED513640

Foster, J., E. Toma, & S. Troske. (2013). Does Teacher Professional Development Improve Math and Science Outcomes and Is It Cost Effective? *Journal of Education Finance, 38*, 255–75.

Fritz, S., M. & MacKinnon, D., P. (2007). Required Sample Size to Detect the Mediated Effect. *Psychol Sci. 18*(3): 233–239. doi:10.1111/j.1467-9280.2007.01882.x.

Fullan, M.& Miles, M. (1992). Getting reform right: What works and what doesn't. *Phi Delta Kappan, 73,* 745–752.

Fullan, M. (1996). Professional Culture and Educational Change. *School Psychology Review*, *25*(4), 496–500.

Fulton, F., Doerr, H., & Britton, T. (2010). STEM teachers in professional learning communities: A knowledge synthesis. *National Commission on Teaching and America's Future*. Retrieved from http://www.wested.org/cs/we/view/rs/1097

Gamoran, A., Secada, W. G., & Marrett, C. B. (2006). The Organizational Context of Teaching and Learning. In M. T. Hallinan (Ed.), *Handbook of the Sociology of Education*, Handbooks of Sociology and Social Research (pp. 37–63). Springer US. Retrieved from http://link.springer.com/chapter/10.1007/0-387-36424-2_3

Garet, M., Porter, A.C., Desimone, L., Birman, B., & Yoon, K.S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*, 915–945.

Gee, J. P. (2004). *Situated Language and Learning: A Critique of Traditional Schooling*.

New York: Routledge.

Gee, J. P. (2005). Semiotic Social Spaces and Affinity Spaces: From The Age of Mythology to Today's Schools. In D. Barton & K. Tusting (Eds.), *Beyond communities of practice: Language, power and social context* (pp. 214–232). Cambridge: Cambridge University Press.

Grigg, J., Kelly, K. A., Gamoran, A., & Borman, G. D. (2013). Effects of Two Scientific Inquiry Professional Development Interventions on Teaching Practice. *Educational Evaluation and Policy Analysis*, *35*(1), 38–56. doi:10.3102/0162373712461851

Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American Psychologist*, *53*(1), 5–26. doi:10.1037/0003-066X.53.1.5

Groome, M., Rankin, J., & Wheary, J. (n.d.). The importance of support & c ollaboration in solving the crisis of stem teacher retention. Retrieved from http://www.demos.org/sites/default/files/publications/STEM-Brief-Demos_NYAS.pdf

Groundwater-Smith, S., & Mockler, N. (2009). *Teacher Professional Learning in an Age of Compliance: Mind the Gap*. Springer.

Guskey, T. & Sparks, D. (2004). Linking professional development to improvements in student learning. In E. Guyton & J. Dangel (Eds.) *Research linking teacher preparation and student performance,* pp. 11-21. Dubuque, Iowa: Kendall/Hunt Publishing.

Halverson, R. (2007). How leaders use artifacts to structure professional community in schools. In L. Stoll & K. S. Louis (Eds.), *Professional learning communities:*

Divergence, depth, and dilemmas (pp. 93–105). Berkshire, UK: Open University Press

Hargreaves, A., & Fullan, M. (2012). *Professional Capital: Transforming Teaching in Every School*. Teachers College Press.

Heller, J. I. (2012). *Effects of Making Sense of SCIENCE[TM] Professional Development on the Achievement of Middle School Students, Including English Language Learners. Final Report. NCEE 2012-4002*. National Center for Education Evaluation and Regional Assistance.

Henstrand, J. L. (2006). Seeking an understanding of school culture: Using theory as a framework for observation and analysis. In V. Anfara & N. Mertz (Eds.), *Theoretical frameworks in qualitative research* (pp.1-22). Thousand Oaks, CA: Sage.

Imai, K., Keele, L., & Yamamoto, T. (2010). Identification, inference, and sensitivity analysis for causal mediation effects. *Statistical Science, 25,* 51–71.

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*,*15*(4), 309-334.

Ingvarson, L., Meiers, M., & Beavis, A. (2005). Factors affecting the impact of professional development programs on teachers' knowledge, practice, student outcomes & efficacy. *Professional Development for Teachers and School Leaders*. Retrieved from http://research.acer.edu.au/professional_dev/1

Jeanpierre, B., Oberhauser, K., & Freeman, C. (2005). Characteristics of professional development that effect change in secondary science teachers' classroom practices. *Journal of Research in Science Teaching*, *42*(6), 668–690. doi:10.1002/tea.20069

Johnson, C. C. (2006). Effective Professional Development and Change in Practice: Barriers Science Teachers Encounter and Implications for Reform. *School Science and Mathematics*, *106*(3), 150–161. doi:10.1111/j.1949-8594.2006.tb18172.x

Johnson, C. & Marx, S. (2009). Transformative professional development: A model for urban science education reform. *Journal of Science Teacher Education, 20(2),* 113-134.

Jr, S. J. G., & Mirkin, C. (2012, June 25). Encouraging STEM Students Is in the National Interest. *The Chronicle of Higher Education*. Retrieved from http://chronicle.com/article/Encouraging-STEM-Students-Is/132425/

Keys, C.W., & Bryan, L.A. (2000). Co-constructing inquirybased science with teachers: Essential research for lasting reform. *Journal of Research in Science Teaching, 38*(6), 631-645.

Kent, A. M. (2004). Improving Teacher Quality through Professional Development. *Education*, *124*(3), 427.

King, K., Shumow, L., & Lietz, S. (2001). Science education in an urban elementary school: Case studies of teacher beliefs and classroom practices. *Science Education, 85,* 89-110.

Knapp, M., S. (2003). Professional development as a policy pathway. *Review of Research in Education*, 27, 109-157.

Lal, S. & Osisioma, I. (2010, May). Classroom perspectives: Observation of the Implementation of a Fourth Grade Immersion Science Inquiry Curriculum. Paper presented at the annual meeting of the *American Educational Research Association Annual Meeting,* Denver, CO.

Leana, C. (2011). *An open letter to Bill and Melinda Gates on the value of social capital in school reform*. Stanford Social Innovation Review.

Lee, V. E., & Smith, J. B. (1996). Collective Responsibility for Learning and Its Effects on Gains in Achievement for Early Secondary School Students. *American Journal of Education*, *104*(2), 103–147. doi:10.2307/1085702

Leithwood, K., Steinbach, R., & Jantzi, D. (2002). School leadership and teachers' motivation to implement accountability policies. *Educational Administration Quarterly, 38*(1), 94-119.

Lewis, E. B. (2011). Secondary Science Teachers' Translation of Professional Development through Affinity-and Institution-identity. Retrieved from http://digitalcommons.unl.edu/teachlearntalks/2/

Lin, F.-L., & Cooney, T. J. (2001). *Making Sense of Mathematics Teacher Education*. Springer.

Little, J.W. (1993). Teachers'professional development in a climate of educational reform. *Educational Evaluation and Policy Analysis, 11*(2), 129-151.

Little, J. W. (2006). Professional community and professional development in the learning-centered school. (Best practices working paper). Washington, D.C.: National Education Association.

Loucks-Horsley, S., & Matsumoto, C. (1999). Research on Professional Development for Teachers of Mathematics and Science: The State of the Scene. *School Science and Mathematics*, *99*(5), 258–271. doi:10.1111/j.1949-8594.1999.tb17484.x

Louis, K. S., & Marks, H. M. (1998). Does Professional Community Affect the Classroom? Teachers' Work and Student Experiences in Restructuring Schools. *American Journal of Education*, *106*(4), 532–575. doi:10.2307/1085627

Louis, K. S., Marks, H. M., & Kruse, S. (1996). Teachers' Professional Community in Restructuring Schools. *American Educational Research Journal*, *33*(4), 757–798. doi:10.3102/00028312033004757

National Research Council. (2000). Inquiry and the national science education standards: A guide for teaching and learning. In S. Olson & S. Loucks-Horsley (Eds.), Committee on the Development of an Addendum to the National Science Education Standards on Scientific Inquiry. Washington, DC: National Academy Press.

National Science Teachers Association (2004). *NSTA Position Statement: Scientific Inquiry*.

Newman, L., Wagner, M., Huang, T., Shaver, D., Knokey, A.-M., Yu, J., Contreras, E., et al. (2011). *Secondary School Programs and Performance of Students with Disabilities: A Special Topic Report of Findings from the National Longitudinal Transition Study-2 (NLTS2). NCSER 2012-3000*. National Center for Special Education Research. 400 Maryland Avenue SW, Washington, DC 20202.

Penuel, W. R., Fishman, B. J., Gallagher, L. P., Korbak, C., & Lopez-Prado, B. (2008). The mediating role of coherence in curriculum implementation. In *Proceedings of the 8th International Conference of the Learning Sciences.* Utrecht, the Netherlands: Erlbaum.

President's Council of Advisors on Science and Technology (2010). Report to the president, Prepare and Inspire: K-12 education in science, technology, engineering, and mathematics (STEM) for America's future. Retrieved from h ttp://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-stemed-report.pdf

Scribner, J. P., Cockrell, K. S., Cockrell, D. H., & Valentine, J. W. (1999). Creating professional communities in schools through organizational learning: An evaluation of a school improvement process. *Educational Administration Quarterly*, *35*(1), 130–160.

Smith, T. M. et al. (2007). Inquiry-oriented instruction in science: Who teaches that way? *Educational Evaluation and Policy Analysis, 29*(3), 169-199.

Spillane, James P., Reiser, Brian J. & Gomez, Louis M. (2006). *Policy implementation and cognition: The role of human, social, and distributed cognition in framing policy implementation.* In Meredith I. Honig (Ed.), New directions in education policy implementation: Confronting policy. (pp. 47-64). State University of New York: Albany.

Sunal, D. and Wright, E. (2006). Teacher perceptions of science standards in K-12 classrooms: An Alabama case study. In D. Sunal & E. Wright (Eds.), *The impact of state and national standards on K-12 teaching* (pp. 123-151). Charlotte, NC: Information Age Publishing.

Supovitz, J. A., & Turner, H. M. (2000). The effects of professional development on science teaching practices and classroom culture. *Journal of Research in Science Teaching, 37*(9), 963–980.

van Driel, J. H., Beijaard, D., & Verloop, N. (2001). Professional development and r eform in science education: The role of teachers' practical knowledge. *Journal of Research in Science Teaching*, *38*(2), 137–158. doi:10.1002/1098-2736(200102)38:2<137::AID-TEA1001>3.0.CO;2-U

Wallace, M. (2009). Making sense of the links: Professional development, teacher practices, and student achievement. *Teachers College Record, 111(2),* 573–596.

Wallace, C. S. (2012). Authoritarian science curriculum standards as barriers to teaching and learning: An interpretation of personal experience. *Science Education*, *96*(2), 291–310. doi:10.1002/sce.20470

Weick, K. E. (1982). Administering Education in Loosely Coupled Schools. *The Phi Delta Kappan*, *63*(10), 673–676. doi:10.2307/20386508

Weick, K. E. (1976). Educational organizations as loosely coupled system. *Administrative Science Quarterly, 21,* 1-19.

Weiss, I. R., Banilower, E. R., McMahon, K. C., & Smith, P. S. (2001). Report of the 2000 national survey of science and mathematics (status report). Chapel Hill, NC: Horizon Research.

White House (2012, February 07). *President Obama to Host White House Science Fair.* Retrieved from http://www.whitehouse.gov/the-press-office/2012/02/07/president-obama-host-white-house-science-fair