Utah State University
DigitalCommons@USU

All Graduate Theses and Dissertations | Graduate Studies

2012

Effects of Sustained Teacher Professional Development on the Classroom Science Instruction of Elementary School Teachers

Nancy Hauck
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/etd

Part of the Teacher Education and Professional Development Commons

Recommended Citation
Hauck, Nancy, "Effects of Sustained Teacher Professional Development on the Classroom Science Instruction of Elementary School Teachers" (2012). *All Graduate Theses and Dissertations*. 1343.
https://digitalcommons.usu.edu/etd/1343

This Dissertation is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact dylan.burns@usu.edu.

EFFECTS OF SUSTAINED TEACHER PROFESSIONAL DEVELOPMENT ON THE

CLASSROOM SCIENCE INSTRUCTION OF ELEMENTARY SCHOOL TEACHERS

by

Nancy Hauck

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

of

DOCTOR OF PHILOSOPHY

in

Education

Approved:

Dr. Todd Campbell
Committee Co-Chair

Dr. Kimberly Lott
Committee Co-Chair

Dr. Suzann Broughton
Committee Member

Dr. Ann M. Berghout Austin
Committee Member

Dr. Deborah Byrnes
Committee Member

Dr. Susan Turner
Committee Member

Dr. Mark R. McLellan
Vice President for Research and
Dean of the School of Graduate Studies

UTAH STATE UNIVERSTIY
Logan, Utah

2012

Copyright © Nancy Hauck 2012

All Rights Reserved

ABSTRACT

Effects of Sustained Teacher Professional Development on the

Classroom Science Instruction of Elementary School Teachers

by

Nancy Hauck, Doctor of Philosophy

Utah State University, 2012

Major Professors: Todd Campbell, Ph.D., & Kimberly Lott, Ph.D.
Department: Teacher Education and Leadership

The purpose of this study was to determine the extent to which sustained teacher professional development in science education affects the classroom instruction of elementary school teachers in third through sixth grade over a 3-year period. The teachers in the study were all elementary endorsed and prepared to be generalists in the content areas.

Science reform has led to more content-specific science standards that are difficult for most elementary teachers to address without professional development. Recent studies on improving elementary science instruction suggest the need for professional development to be long term, embedded in teaching practice in the classroom, and rooted in research on how children learn science. The researcher examined changes in classroom instruction over a 3-year period of teachers who participated in a professional development program designed to meet the elementary science education reform based on recommendations from the National Research Council's report, *Taking Science to School: Learning and Teaching Science in Grades K-8*.

The data that were analyzed to determine the effects of the professional development came from classroom observations of two sets of teachers, one of which was the control set (n = 20). The other was the experimental set (n =22). Classroom observations were administered one time each year over 3 years of treatment to determine whether sustained professional development in science impacted teacher practices in the classroom.

This study suggested that classroom science instruction did significantly change through sustained professional development intervention. It also suggested that teaching practices improved in the areas of talk and argument, investigation and inquiry, modeling and representations, alignment with science core concepts, and addressing science misconceptions. Furthermore, findings indicated that teachers who received sustained professional development were more likely to have higher overall effective science instruction scores.

(210 pages)

PUBLIC ABSTRACT

Effects of Sustained Teacher Professional Development on the

Classroom Science Instruction of Elementary School Teachers

by

Nancy Hauck, Doctor of Philosophy

Utah State University, 2012

The purpose of this study was to determine the extent to which sustained teacher professional development in science education affects the classroom instruction of elementary school teachers in third through sixth grade over a 3-year period. The teachers in the study were all elementary endorsed and prepared to be generalists in the content areas.

Science reform has led to more content-specific science standards that are difficult for most elementary teachers to address without professional development. Recent studies on improving elementary science instruction suggest the need for professional development to be long term, embedded in teaching practice in the classroom, and rooted in research on how children learn science. The researcher examined changes in classroom instruction over a 3-year period of teachers who participated in a professional development program designed to meet the elementary science education reform.

Classroom observations were administered to determine whether sustained professional development in science impacted teacher practices in the classroom. Observations were made of two groups of teachers, one of which was the control group of 20 teachers and the other was the experimental group of 22 teachers. The control group was used to illustrate nontreatment and was observed once during year one of the study. The treatment group was observed three times, one time each year over 3 years of treatment. Two observation instruments were used to evaluate classroom science instruction. The Summary Judgment of Science Instruction was used to evaluate overall science instruction. The PESTL Observation Protocol was used to evaluate five components of reformed science instruction, which included: talk and argument, investigation, modeling, content alignment, and addressing misconceptions. The data were analyzed to determine the effects of the professional development on classroom instruction overtime.

This study suggested that classroom science instruction did significantly change through sustained professional development of 3 years. Findings indicated that 1 year of professional development does not appear to significantly improve overall science instruction or the practices of reformed science instruction of elementary teachers who participate when compared to nonparticipants. Findings indicated that 2 years of sustained professional development appears to significantly improve the practices of talk and argument, investigation and content alignment of elementary teachers who participate when compared to 1 year of treatment. Furthermore, findings indicated that 3 years of sustained professional development appears to concomitantly improve overall science instruction and all five components of reformed science instruction of elementary teachers who participate when compared to 1 and 2 years of treatment. Notably, overall science instruction and the components of modeling, and addressing misconceptions significantly improved only after 3 years of professional development. This study suggested that 3 years or more of professional development is necessary to significantly improve overall science instruction and all five components of reformed science instruction concomitantly.

DEDICATION

To my husband, Jason, and our children Amanda & Shaye,

Jeffrey, Stephanie, Bradley, and Amy.

And to Dixie State College for instilling the love of learning in my heart,

and for allowing me to teach it to others,

who in turn will pass the love of learning to the next generation.

ACKNOWLEDGMENTS

The process of writing this dissertation, like most meaningful endeavors, was filled with extraordinary experiences, memories, and people. First, I wish to thank the co-chairs of my committee, Drs. Todd Campbell and Kimberly Lott. Their thoughtful guidance and support were helpful to me as a novice researcher. Dr. Campbell's inspired direction early on set me on the right course and his kind support and enthusiasm have truly uplifted me throughout this study. Dr. Lott's promptings to submit my study to journals and conferences have contributed much to my dissertation experience.

My dissertation committee, Dr. Ann M. Berghout Austin, Dr. Suzanne H. Broughton, Dr. Deborah Byrnes, and Dr. Susan Turner, provided meaningful feedback, which improved the direction and quality of this dissertation. Specifically, I would like to thank Dr. Broughton, the statistical expert on my committee, for her tireless work. Dr. Broughton taught me in two quantitative research methods courses and spent many hours working with me and getting answers to my questions. She was also a valuable resource on my dissertation topic of elementary science education.

I wish to thank Brett Moulding, director of Partnerships in Effective Science Teaching and Learning (PESTL), for the wonderful opportunity he provided for me to study this professional development program. He allowed me to objectively examine the program and data sets gathered in the classroom observations. Brett facilitated my analysis of the program's effectiveness and welcomed my suggestions of how to improve the PESTL program and data collection process. I would also like to thank the teachers who formed the treatment and control groups of this study. They are the unseen faces in this work.

My doctoral experience was greatly enhanced by my professors and cohort members. Over the past 4 years I have learned much from the outstanding USU TEAL faculty members who have taught me. Mostly I have grown to love my fellow cohort members in the 2008 Distance Doctoral Program. They are an amazing bunch of “cats” that USU “herded” together.

Next, I wish to thank Dixie State College for instilling the love of learning in my heart, and for allowing me to teach it to others. Dixie is my home and I have been blessed to have the red sand of Dixie in my shoes. I consider it an honor to be part of the faculty in the Department of Education at Dixie State College. One of the greatest joys in my life is to educate future teachers, who will in turn pass the love of learning to the next generation.

Finally, those who have meant the most to me in this effort have been my family. Many thanks to my sister- and bother-in-law, Jenny and Todd Condie, and their children for opening their home to me when I spent the summer months in Logan. They have a special place in my heart for that. My large extended family—on both sides—has sent encouraging words and support as well. I thank my children, Amanda and Shaye, Jeffrey, Stephanie, Bradley, and Amy. They have been patient and understanding of their busy mom. In closing, it is my husband, Jason, who deserves the most credit. He has been my greatest supporter.

Nancy Hauck

CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND BACKGROUND

Amy King has taught fourth grade at Canyon Elementary for 7 years. Over the past 3 years, as grade-level team leader, she has been instrumental in helping her school meet the No Child Left Behind Act's adequate yearly progress (AYP) in language arts and math. However, during the past 2 years, the school's overall science achievement scores have been low. The school administrator invited Ms. King to participate in science professional development with the other third- through-sixth grade teachers at the school. During the first week's session of professional development, Ms. King discovered she had many science misconceptions. She realized that her science understanding needed to be transformed in order for her to teach science accurately to her students. The professional development team told Ms. King about new research on how children learn science and presented recommended teaching practices for science education reform. After the first week of professional development, Ms. King realized that she still had much to learn about teaching science in her elementary classroom.

The vignette above and those used throughout this paper are examples of the collective experiences that teachers have in elementary science professional development. The teacher's name is a pseudonym and not an actual teacher in the study.

Background and Significance of the Study

For the last two decades the science competency scores for students in the U.S., when compared with those of other developed nations, have been low (International Association for the Evaluation of Educational Achievement [IEA], 1995, 2003). Results of the Trends in International Mathematics and Science Study (TIMSS), involving a half-million students from 41 countries, convincingly illustrate that American fourth graders performed poorly, middle school students performed even more poorly, and high school students performed worst of all (Forgione, 2006; IEA, 2003). The science tests indicated that U.S. fourth grade students ranked in 12th position, which was the 50th percentile out of 26 industrialized nations. U.S. eighth-grade students scored in the bottom third, and by twelfth-grade, U.S. students finished nearly last (Forgione, 2006). After examining these findings, Forgione (2006) concluded, “By the time our students are ready to leave high school—ready to enter higher education and the labor force—they are doing so badly in science that they are significantly weaker than their peers in other countries” (p. 2). These findings support the need for U.S. science education reform.

Furthermore, the National Center for Education Statistics (NCES, 2009) reported that only 34% of U.S. fourth graders, 30% of eighth graders, and 21% of twelfth graders performed at or above the National Assessment of Educational Progress (NAEP, 2009) proficient level in science. Only students reaching proficient level demonstrated competency over challenging science subject matter. In the state of Utah, more specifically, the setting for this study, 37% of fourth graders and 37% of eighth graders performed at or above the NAEP proficient level in science (NCES, 2009). Even though the state of Utah performed higher than the national average, there is still much room for improvement.

Most elementary teachers are not sufficiently prepared to teach science effectively in terms of both knowledge of science content and familiarity with inquiry-based science instruction, which impedes efforts to improve students’ science achievement (Kennedy, 1998; Loucks-Horsley, Love, Stiles, Mundry, & Hewson, 2003; NCES, 2009). The literature is replete with examples indicating that elementary teachers are often not adequately prepared in science and not comfortable teaching science (Bruning, Schraw, & Ronning, 1999; Hiebert, 1997; Kennedy, 1998; Loucks-Horsley, 1998; Loucks-

Horsley et al., 2003; Weiss, 1994; Yager, 2000). After examining the findings of TIMSS 1995 report, Stigler and Hiebert (1999) concluded, "American teachers aren't incompetent, but the methods they use are severely limited, and American teaching has no system in place for getting better. It is teaching, not teachers, that must be changed" (p. 10). These findings support the need for innovative and transformational professional development in elementary science education.

Recent standards documents and research on improving elementary science learning build on a model of instruction that is starkly different from current teaching practice (Akerson, 2005; Hiebert, 1997; National Research Council [NRC], 2005, 2007; Schneider & Krajcik, 2002; Thorson, 2002). Current practice falls short in that most elementary science instruction simply relies on telling students what scientists have already discovered or in asking students to follow the steps of the "scientific method" (NRC, 2005). Innovative teaching practice "helps students develop the knowledge, skills, and attitudes that will enable them to understand what it means to "do science" and to participate in the larger science community" (NRC, 2005, p. 398). However, many elementary level teachers feel unprepared to implement the innovative teaching practices prescribed for science education reform (Akerson, 2005; Birman, Desimore, Porter, & Garet, 2000; Duschl, Schweingruber & Shouse, 2006; Kennedy, 1998; Loucks-Horsley et al., 2003; NCES, 2009).

Most elementary teachers are trained to be generalists in the content areas and are strongly prepared in pedagogical practices, reading skills, basic language arts, and mathematics content areas (Birman et al., 2000). Science reform has led to more content-specific science standards that are difficult for most elementary teachers to address without additional education (Birman et al., 2000; Duschl et al., 2006; Kennedy, 1998; Loucks-Horsley et al., 2003). Reformed elementary science standards cover content-specific objectives in the areas of life science, space and earth science, physical science, engineering, and technology (NRC, 2011). Few elementary teachers have sufficient background in these areas to teach the learning objectives to their students.

The limited time that is available for building science content literacy and teaching methods in preservice elementary programs is one factor that leads teachers to feel unprepared to teach science. The NCES (Parsad, Lewis, & Farris, 2001) reported that most elementary teacher preparation programs require the minimum general education science coursework and only one elementary science methods. Also, veteran teachers in the workforce may be far removed from their preservice days, making it more likely that the focus of current science reforms represent significant changes from when they participated in their teacher education programs (NCES, 2001).

At the elementary level, where most teachers are assigned to teach science and other academic subjects to one group of students, 76% of teachers reported feeling very well-qualified to teach reading and roughly 60% felt very well qualified to teach mathematics and social studies. In contrast, only 28% felt very well qualified to teach life science; and fewer than 10% felt very well qualified in the physical sciences (NCES, 2009; Weiss, 1994). This perceived lack of preparation is often connected to the limited amount of time devoted to the elementary science curriculum (Akerson, 2005; Enoch & Riggs, 1990; Fitch & Fisher, 1979; NCES, 2009; Riggs, 1991; Weiss, 1994). What's more, teachers with little science preparation may develop a negative attitude toward science and therefore avoid its teaching (Riggs, 1991).

Professional development plays a key role in addressing the gap between teacher preparation and standards-based reform. It can serve as a powerful mechanism to improve instruction and teacher self-efficacy in science education. In a study of 44 inservice elementary teachers, Sottile, Carter, and Murphy (2002) found that teachers who had participated in science professional development improved in self-efficacy beliefs. At the end of the science professional development, the inservice teachers thought they could better motivate students to enjoy science, and felt more competent to answer questions about science experiments. These teachers believed they could better plan science lessons using constructivist techniques and felt more competent in their own science understandings. Furthermore, these teachers thought they could better assist their colleagues in planning and teaching with science instruction.

The wave of reform that has swept across the U.S. over the last two decades has created a climate of change that requires school districts, administrators, schools, and classroom teachers to reexamine their core beliefs regarding science teaching and learning. These reform efforts in one way or another require systemic transformational change. To implement improvements in science education, reform efforts connect science teachers with researchers and practitioners of current findings and projects in science education. Two major national policy initiatives, Project 2061 (Rutherford & Ahlgren, 1990) and the National Science Education Standards (NSES; NRC, 1996) lead the efforts to reconstruct science education.

Project 2061 (Rutherford & Ahlgren, 1990) was a long-term initiative of the American Association for the Advancement of Science (AAAS) to help all Americans become literate in science, mathematics, and technology. To achieve that goal, Project 2061 conducted research and developed tools and services that educators, researchers, and policymakers used to make critical and lasting improvements in the nation's education system. The project's areas of expertise include learning goals and curriculum, assessment, and teacher development.

The NSES (NRC, 1996), an initiative of the NRC, presented a vision of a scientifically literate populace. The standards outlined what students need to know, understand, and be able to do to be scientifically literate at different grade levels (NRC, 1996). They described "an educational system in which all students demonstrate high levels of performance, in which teachers are empowered to make the decisions essential for effective learning, in which interlocking communities of teachers and students are focused on learning science, and in which supportive educational programs and systems nurture achievement" (NRC, 1996, p. 2). The organization of the NSES included standards for science teaching, professional development for teachers of science, assessment in science education, science content, and science education systems. For the vision of science education described in the NSES to be attained, the standards contained in all six areas need to be implemented (NRC, 1996).

Very recently, the NRC (2011) released a new framework for the NSES, *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. The foreword of this standard document stated:

> This project capitalizes on a major opportunity that exists at this moment—a large number of states are adopting common standards in mathematics and English/language arts and thus are poised to consider adoption of common standards in K-12 science education. The impetus for this project grew from the recognition that, although the existing national documents on science content for K-12 (developed in the early to mid 1990s) were an important step in strengthening science education, there is much room for improvement. (NCR, 2011, p. viii)

This framework represents the first step in a process to create new standards in K-12 science education and will likely impact future reform efforts in science education.

Both Project 2061 and NSES emphasized teacher professional development as the crucial cornerstone in reforming science education. In fact, many education scholars believe that a critical component of any educational reform effort should be to provide teachers with opportunities and appropriate support structures that encourage the significant work of ongoing development of pedagogical practice (Darling-Hammond & McLaughlin, 1995; Garet, Porter, Desimone, Birman, & Yoon, 2001; Sparks & Hirsh, 1997). However, much of the professional development offered to teachers simply does not meet the challenges of the reform movement (Darling-Hammond & Baratz-Snowden, 2007; Elmore, 2002; Stein, Smith, & Silver, 1999). In practice, professional development covers a vast array of specific activities, everything from highly targeted work with teachers, focusing on specific curricula, to short "hit-and-run" workshops designed to familiarize teachers and administrators with new ideas or new requirements (Elmore, 2002). To implement instruction that is consistent with the research base, teachers need substantial, ongoing, and systematic support for their own learning (NRC, 2007, 2011; Neuman, 2009).

The Problem

Nationally, a concentrated effort has been made to reform science education through teacher professional development; however, most of the programs are short term and fail to offer teachers the science literacy and continued support they need in order to transform the science teaching and learning in their classrooms (Birman et al., 2000; Duschl et al., 2006; Elmore, 2002). Recent studies on improving science instruction suggest the need for professional development to be long term, embedded in teaching practice in the classroom, and rooted in research on how children learn science (Duschl et al., 2006; van Driel, Beijaard, & Verloop, 2001; Weiss, Pasley, Smith, Banilower, & Heck, 2003; Yager, 2000). The NRC (2007) reported that innovative approaches to professional development "should be rooted in the science that teachers teach and should include opportunities to learn about science, about current research on how children learn science, and about how to teach science" (p. 7). These recommendations suggest a need for professional development to be sustained over a number of years and based on elementary science content and effective teaching practice.

To meet the call for reform in science education, the NRC's report *Taking Science to School: Learning and Teaching Science in Grades K-8* compiled several recommendations for effective professional development (Duschl et al., 2006). This report specifically calls for professional development to be: (a) long term, (b) to focus on building teacher science literacy, and (c) to focus on reforming science teaching practice. By long term, the report suggested that professional development be sustained for three or more years. The NRC (1996) defined science literacy as the "knowledge and understanding of scientific concepts and processes required for personal decision making, participation in civic and cultural affairs, and economic productivity" (p. 22). The reform instructional practices outlined in the report include, making students' learning visible through talk and argument, models, and investigations; aligning instruction to science core concepts, and appropriately addressing science misconceptions (NRC, 2007).

Professional developers across the U.S. are implementing such programs (Moulding & Baird, 2008; NCES, 2009; NRC, 2011). However, do these professional development programs actually transform science instruction in the elementary classroom? There have been relatively few reports of quantitative studies investigating whether these new long-term teacher professional development programs in elementary science education actually yield changes in instructional practices over time. These programs need to be studied to determine their success in reforming science education. This study endeavored to examine such a program and its potential effectiveness in transforming classroom science instruction over time.

This paper outlines a quantitative study of a sustained professional development program for reforming elementary science education. The following sections will reveal the purpose and objectives of the study, the research questions, methodological strategies and assumptions, rationale for the study, and definition of terms.

Purpose and Objectives

The purpose of this study was to determine the extent to which sustained teacher development in elementary science education affects classroom instruction over a 3-year time period. There were three main objectives of this study.

1. To determine whether long-term teacher development in elementary science education actually yields changes in instructional practices

2. To determine the areas of instruction that are most influenced by long-term teacher development in elementary science education

3. To determine how instruction changes and if patterns of change exist over a 3-year time period with teachers who participate in sustained teacher development in elementary science education

Professional development, designed to meet the objectives of educational reform, intends to transform classroom practices to improve student learning and meet the changing emphasis of science education. This study examined the effects of a particular professional development program on the classroom instruction of participating elementary school teachers. Changes in student achievement scores were not within the scope of this study but would provide an excellent focus for a follow-up study. Transformation of elementary science classroom *instruction* was the focus of this study.

Research Questions

The central question of this study was, does sustained teacher professional development, which was designed to meet recommendations for reform in elementary science education, transform classroom instruction over a 3-year time period? In relation to this question, three questions, with related subquestions, were explored.

1. Does sustained professional development of 3 years, which was designed to

meet recommendations for reform in science education, transform overall classroom science instruction of elementary school teachers?

2. Does sustained professional development of 3 years transform the instructional practice of elementary school teachers in specific components of reformed science education?

- 2a. Does sustained professional development affect the practice of classroom discussion and argumentation in the science instruction of elementary school teachers?
- 2b. Does sustained professional development affect the practice of scientific investigation in the classroom science instruction of elementary school teachers?
- 2c. Does sustained professional development affect the practice of teacher and student modeling in the classroom science instruction of elementary school teachers?
- 2d. Does sustained professional development affect the practice of aligning science instruction to core standards in the classroom instruction of elementary school teachers?
- 2e. Does sustained professional development affect the practice of appropriately addressing student science misconceptions in the classroom instruction of elementary school teachers?

3. How does sustained professional development of 3 years transform overall classroom science instruction and the specific components of reformed science

instruction of elementary school teachers over time?

3a. What components of reformed science instruction are most influenced by sustained professional development over time?

3b. What components of reformed science instruction are least influenced by sustained professional development over time?

3c. Are there patterns of change among the components of reformed science instruction of elementary school teachers over time?

3d. Are there interactions between performance in overall classroom science instruction and the practice of specific components of reformed science instruction of elementary school teachers over time?

Methodological Strategies and Assumptions

The study examined the effects of sustained professional development on the classroom science instruction of elementary school teachers. The study connected to the paradigm of educational reform, specifically in elementary science education. The study fell within the social constructivist tradition with the transformative learning theory serving as a theoretical lens. Social constructivism is a sociological theory of knowledge that applies the general philosophical base of constructivism into social settings, wherein groups construct knowledge with one another, collaboratively creating a culture of shared artifacts with shared meanings. When one is immersed within a culture of this sort, one is learning all the time about how to be a part of that culture on many levels. Social constructivism is largely attributed to Lev Vygotsky (Kukla, 2000). The transformational learning theory, articulated by Mezirow (2000), referred to the process by which adults transform their taken-for-granted frames of reference to make them more "inclusive, discriminating, open, emotionally capable of change, and reflective" (pp. 7-8). This transformation allows adult learners to generate ideas and opinions that are more accurate and defensible to guide action (Mezirow, 2000). Furthermore, the theory is described as being social constructivist in nature and oriented toward explaining the way adult learners interpret and reinterpret their experiences (Mezirow, 1991).

The research methodological strategies of this study followed a quantitative orientation with a quasi-experimental design. The researcher studied two groups: a treatment group consisting of teachers who participated in the professional development program, and a control group consisting of teachers who did not participate in the professional development program. The axiological assumption was that the goal of the researcher was to remain objective and value free. The ontological assumption was that changes in elementary science instruction would be observable and measurable; therefore, information and data were gathered through objective and quantifiable measures.

Rationale for the Study

To meet the current recommendations of leading experts and researchers, sustained teacher development programs have been designed to reform elementary science education (Minuskin, 2009; Moulding & Baird, 2008; NRC, 2007; Ornek, 2008; Richardson & Liang, 2008). Empirical studies are necessary to assess the effectiveness of these new professional development programs. Educational experts, policy makers, grant awarding agencies, and politicians rely on observable, objective, and quantifiable measures to make decisions about a new professional development program's success. This study aimed to evaluate the effectiveness of such a program by quantitatively examining instructional changes in teachers who participated in sustained elementary science professional development.

Definition of Terms

The researcher has defined the following terms according to their use in this dissertation study.

Professional development: According to the Association for Supervision and Curriculum Development (2007),

> this term refers to study that is directly focused on helping to achieve student learning goals and supporting student learning needs and is a collaborative endeavor—teachers and administrators working together in planning and implementation. It is school-based and job-embedded, is a long-term commitment, is differentiated, and is tied to district goals. (p. 2)

Science education reform: Science education reform examines strategies for implementing improvements in science education by connecting science teachers with expert researchers and practitioners of current findings and projects in science education. Project 2061 (Rutherford & Ahlgren, 1990) and the NSES (NRC, 1996) lead the movement to reconstruct science education. Science education reform pushes to connect science learning to the practices of scientists and to teaching science as inquiry instead of the traditional didactic approach.

Best practice: A best practice is a technique or methodology that, through experience and research, has proven reliably to lead to a desired result. A commitment to using the best practices in any field is a commitment to using all the knowledge and technology at one's disposal to ensure success (Zemelman, Daniels, & Hyde, 2005).

Partnership for Effective Science Teaching and Learning (PESTL): PESTL is a 3-year professional development program funded by a grant from the Utah State Office of Education under the U.S. Department of Education's ESES Title II part B Mathematics and Science Partnership grant program. PESTL seeks to improve student learning through sustained teacher professional growth and science literacy.

Professional learning communities (PLCs): Although there is no universal definition of a PLC, DuFour, Eaker, and Many (2006) defined it as "educators committed to working collaboratively in ongoing processes of collective inquiry and action research to achieve better results for the students they serve" (p. 21).

Four strands of science learning: These "strands" encompass the knowledge and reasoning skills that students eventually must acquire to be considered proficient in science as established by the NRC (2007). The four strands include: (a) understanding science explanations, (b) generating scientific evidence, (c) reflecting on scientific knowledge, and (d) participating productively in science (NRC, 2007).

This chapter has given the background and significance of a study examining the effects of teacher professional development on the science instruction of elementary classroom teachers. A problem statement was given followed by the purpose and objectives of this study. The research questions, methodological strategies and assumptions, rationale for the study, and definition of terms were provided to further outline the study. The next chapter will provide a review of literature relating to the theoretical lens of this study, professional development in science, effective instructional models and practices in elementary science, and recent research on professional development in elementary science education.

CHAPTER 2

REVIEW OF LITERATURE

Ms. King was intrigued with the professional development session that introduced the components of effective instruction for science education reform. She had worried about her students' understanding of science core content and realized that changing her instruction would be the best way to improve learning in her class. Ms. King enjoyed the hands-on science learning she received in the professional development; however, she did not feel confident yet in her ability to teach science concepts in a way that gave her students deep understanding. She acknowledged that this first year of professional development was helping transform her understanding of science teaching and learning. She was gaining science literacy and understanding of best practices but felt she needed more time to actually incorporate these concepts and methods into her classroom instruction. Ms. King was pleased that the professional development program would be sustained over 3 year. She trusted that over time she would be more effective in implementing the recommended instructional practices that she was learning about in the professional development.

Introduction

Educational reform efforts frequently focus on science education. Specifically, reformers suggest that teachers use inquiry-based, student-centered instructional practices that facilitate students' construction of knowledge (Abd-El-Khalich & Akerson, 2004; Akerson, 2005; Lemke, 2001; Birman et al., 2000; McIntyre & Hagger, 1992; NRC, 1996; Thorson, 2002). Recently released standard documents in elementary science education indicate that the current science curriculum must also be reformed, suggesting that reformed curriculum should include fewer topics that can be taught and studied in greater depth (Berland & McNeill, 2010; Michaels, Shouse, & Schweingruber, 2008; NRC, 2011).

Many states and school districts have made science education a part of their overall effort to improve instruction for students in their schools (Moulding & Baird, 2008; Schneider & Krajcik, 2002). However, reform-based curriculum designed to support students' construction of knowledge in science relies on teachers who possess enhanced science literacy and an understanding of how students learn science (NRC, 1999; National Science Teachers Association, 2003; Thorson, 2002). For many teachers this will mean significant study of science concepts and significant changes in their instructional practices. Since what teachers do in their classrooms depends largely on their knowledge, they will need to learn a great deal to be able to enact reform-based curriculum (NRC, 2007; Schneider & Krajcik, 2002).

This review of literature will connect the current research and reports relating to teacher professional development in elementary science education and the desired outcomes of reformed classroom science instruction of elementary school teachers. A comprehensive review of literature will focus on professional development and recommended practices in elementary science teaching. This review is organized around four questions.

1. Which research tradition and learning theory would serve as a useful theoretical lens in studying professional development for elementary science education?

2. What is the reformed instructional model and what is observable and measureable?

3. Which professional development characteristics and components have proven effective in reforming the classroom science instruction of elementary school teachers?

4. Which recent studies have been conducted on the effect of professional

development on elementary science instruction and how do they relate to this study?

In the first section, social constructivism and transformative learning theory will be discussed as the theoretical lens of this study and connections will be made that illuminate how these theories relate to teacher professional development in elementary science education reform.

Theoretical Framework

The theoretical framework of the professional development being examined in this study falls within the social constructivist tradition, with transformative learning theory serving as a lens. The context for this study also includes educational reform, specifically in elementary science education. Current conceptualizations of social constructivism draw heavily on the work of Lev Vygotsky (1896-1934), who began a tradition in social science termed cultural-historical psychology. Cultural-historical psychology is defined as “the study of the development of psychological functions through social participation in societally-organized practices” (Chaiklin, 2001, p. 21). From this area of cognitive psychology stems the sociocultural learning theory (Chaiklin, 2001).

Vygotsky (1978) studied the importance of learning in social settings and the impact of the assistance of more capable others on the development of the learner. He maintained that children rely on the example and skills of adults and more competent peers to gradually develop abilities to do certain tasks, such as talking about a scientific concept or solving a complex math problem. For Vygotsky and other sociocultural theorists, the social nature of cognitive development is captured in the concept of intersubjectivity, which refers to mutual, shared understanding among participants in an activity (Vygotsky, 1978; Wertsch, 1998).

According to Tharp and Gallimore (1988), this sociocultural perspective has profound implications for teaching, schooling, and education. Vygotsky's work principally centered on children; however, identical processes occur in adult learners. Thus, the influence and interactions between a more knowledgeable individual, such as an expert, peer, or mentor, and an adult learner becomes a critical phenomenon for study in sociocultural learning.

Though developed primarily to explain the teacher-student relationship, Vygotsky's (1978) theory assisted in understanding the process of experienced teachers mentoring novice teachers or professional developers providing continued education and support to practicing teachers. This study of elementary science educational reform and professional development fits comfortably within a social constructivist tradition.

Transformational learning theory, developed by Mezirow (1991, 2000), represents a social constructivist theory that applies primarily to adults. Transformative learning is described as:

> the process by which we transform our taken-for-granted frames of reference (meaning perspectives, habits of mind, mind-sets) to make them more inclusive, discriminating, open, emotionally capable of change, and reflective so that they may generate beliefs and opinions that will prove more true or justified to guide action. (Mezirow, 2000, pp. 7-8)

The theory explained how adult learners interpret and reinterpret their experiences in social settings, wherein groups construct knowledge collaboratively, creating a culture of shared practices and meanings.

Transformational learning theory grew out of Mezirow's research in 1978 on women reentering higher education (Glickman, Gordon, & Ross-Gordon, 2009; Mezirow & Taylor, 2009). Mezirow found that when these women encountered significant life events, such as returning to school because of divorce, they were often faced with a disorienting dilemma resulting in perspective transformation (Mezirow, 2000). Elementary teachers, who engage in learning new concepts in science content knowledge and teaching practice, often experience a similar disorienting dilemma and perspective transformation. In collaborative professional development settings these teachers may construct new knowledge and reform their teaching practices.

Transformational learning theory embraces two basic kinds of learning modes: instrumental and communicative (Mezirow & Taylor, 2009). Instrumental learning focuses on learning through task-oriented problem solving and cause-effect relationships. Instrumental learning may involve controlling the environment or managing other people. In the instrumental mode understandings are validated by empirical evidence to ascertain the truth of a belief, association, concept, value, feeling, or world-view (Mezirow & Taylor, 2009).

Communicative learning focuses on how individuals communicate their feelings, needs, and desires (Taylor, 2000). We validate or justify our contested belief, association, concept, value, feeling, or world view through dialogic discourse. This dialogue facilitates social constructivist process in which adults collaborate to develop understanding and meaning. Discursive assessment is the type of dialogue we participate in with others whom we feel to be informed, objective, and rational. We view these others as more competent and capable to evaluate our problematic understandings (Mezirow & Taylor, 2009). Adults learn best from more competent peers while engaged in both instrumental and communicative modes (Glickman et al., 2009; Mezirow & Taylor, 2009).

Synthesis studies report that reform in science teaching is best accomplished by viewing teacher development through a transformative lens (NRC, 2011; Stein et al., 1999). Transformative learning has practical application for preservice teacher education and inservice professional development programs. Teacher professional development programs in elementary science education should draw upon transformational learning theory to develop circumstances that encourage teachers to move toward a frame of reference that is more comprehensive, self-reflective, and integrative of experience. These professional development programs should address misconceptions about science and science education through providing transformational learning experiences for participating teachers.

The transformational learning theory was selected as the lens for this study for two compelling reasons. First, the professional development program examined in this study was designed to align with transformational learning theory concepts. It was designed to meet the needs of inservice elementary teachers by address misconceptions about science and science education. The program, which was designed to meet recommendations for reform in science education, provides sustained transformational learning experiences for teachers. Second, the research design of this study is to examine the transformation of classroom science instruction of elementary school teachers, who participated in this sustained professional development of 3 years. Transformational learning theory posits that adult learning takes time (Mezirow, 2000) and this study set forth to examine patterns of change in adult science learning and teaching over time. The following section outlines the reformed instructional models and the observable practices of effective elementary science instruction that will be examined in this study.

Reformed Instructional Models and Practices of Effective Teaching

Science Literacy

Throughout the history of science education, no definitions for science literacy have been agreed upon; thus, no generally accepted basis for establishing policy, research, curriculum, and teaching regarding science literacy exists (Hassard & Dias, 2008; Hodson, 2008). However, meeting a national goal to reform science education to improve scientific literacy depends upon a single, consistent definition. The term "scientific literacy" first appeared in the educational literature in papers by Paul Hurd (1958) and Richard McCurdy (1958). It was eagerly accepted by others as a valuable concept, but it had little in the way of precise or agreed connotation until Pella, O'Hearn, and Gale (1966) suggested that scientific literacy comprises an "understanding of the basic concepts of science, the nature of science, the ethics that control scientists in their work, and the interrelationships of science, technology and society" (p. 199). Science for All Americans (American Association for the Advancement of Science [AAAS], 1989) drew upon very similar categories to define a scientifically literate person as "one who is aware that science, mathematics, and technology are interdependent human enterprises with strengths and limitations; understands key concepts and principles of science; is familiar with the natural world and recognizes both its diversity and unity; and uses scientific knowledge and scientific ways of thinking for individual and social purposes" (p. 4).

Hodson (2008) argued that scientific literacy for active citizenship, responsible environmental behavior and social reconstruction lies more in learning about science than it does in learning science. Hodson stated that

> ...we should place considerably more emphasis on those elements of the history, philosophy and sociology of science that would enable students to leave school with a robust knowledge about the nature of scientific inquiry and theory building, an understanding of the role and status of scientific knowledge, an ability to understand and to use the language of science, some insight into the sociocultural, economic and political factors that impact the priorities and conduct of science, and some experience of conducting authentic scientific investigation. (p. 20)

While Hodson did not dismiss the importance of knowledge of the major concepts, ideas and theories of science, he maintains that if students acquire good learning habits and attitudes about science in the school years, it will be easier for them to gain additional science knowledge later on.

In *Science Matters* (2009), Hazen and Trefil stated that scientific literacy "constitutes the knowledge you need to understand public issues" (p. xii). They further clarified by explaining that it is a mix of facts, vocabulary, concepts, history and philosophy. In this sense scientific literacy is not the specialized knowledge and skills of the experts, rather it is the kind of knowledge used to understand news and political discourse of the day as it relates to science (Hazen & Trefil, 2009).

The NSES (NRC, 1996) supported the concept of science literacy and have defined it as the "knowledge and understanding of scientific concepts and processes required for personal decision making, participation in civic and cultural affairs, and economic productivity" (p. 22). As this definition indicates, the NSES standards promote knowledge of scientific concepts as well as understanding of the science process as essential and equally important features of science literacy (NRC, 1996). However, in light of past and more recent definitions of scientific literacy, a more complete definition of the term would include the history of science, nature of science and language of science.

For this study, the following definition will be employed: Scientific literacy is the (a) understanding of key scientific concepts, (b) ability to understand and use the processes and language of science, (c) knowledge of the nature of science including its strengths and limitations, (d) knowledge of the history and philosophy of science and the interrelationships of science, technology and society, and (e) ability to use scientific knowledge and scientific ways of thinking for personal decision making, participation in civic and environmental matters and economic productivity. This definition of scientific literacy meets the call of science education reform and draws on the NRC's (1996) definition with additional details from AAAS (1989), Hazen and Trefil (2009), Hodson (2008), and Pella and colleagues (1966).

The NSES outlined a new vision of elementary science education through promoting changing emphases in science content and assessment of science knowledge and understanding. These initiatives and standards promote science literacy and educational reform. The NSES changing emphasis will be the focus of the next section.

NSES for Scientific Content and Process

The NSES view science education as something that students "do" rather than something that is "done to them." The NSES strongly emphasize integrating the processes and nature of science with content knowledge in the various scientific disciplines as a student progresses through the elementary grades (Labov, 2006; NRC, 1996). The NSES promotes a very different way of presenting content and assessing students' knowledge of science (NRC, 1996). Table 2.1 outlines the changing emphasis on scientific content and process. Table 2.2 outlines the changing emphasis on assessment of scientific knowledge and understanding.

As illustrated in Tables 2.1 and 2.2, the NSES stress that not all content is of equal importance; therefore, the standards recommend that elementary teachers cover fewer

Table 2.1

The NSES Changing Emphasis on Scientific Content and Process

Less emphasis on…	More emphasis on…
Knowing scientific facts and information	Understanding science processes and developing abilities of inquiry
Studying subject matter disciplines (e.g., physics, earth sciences) for their own sake	Learning subject matter discipline in the context of inquiry, technology, science in personal and social perspectives, and history and nature of science
Separating science knowledge and science process	Integrating all aspects of science concepts
Covering many science topics	Studying a few fundamental concepts
Implementing inquiry as a set of processes	Implementing inquiry as instructional strategies, abilities, and ideas to be learned

Note. From *National science education standards* (p. 113), by National Research Council, National Committee on Science Education Standards and Assessment, 1996, Washington, DC: National Academy Press.

Table 2.2

The NSES Changing Emphasis on Assessment of Scientific Knowledge and Understanding

Less emphasis on…	More emphasis on…
Assessing what is easily measured	Assessing what is most highly valued
Assessing discrete knowledge	Assessing rich, well-structured knowledge
Assessing scientific knowledge	Assessing scientific understanding and reasoning
Assessing to learn what students do not know	Assessing to learn what students do understand
Assessing only achievement	Assessing achievement and opportunity to learn
End-of-term assessment by teachers	Students engage in ongoing assessment of their work and that of others
Development of external assessments by measurement experts alone	Teachers involved in the development of external assessments

Note. From *National science education standards* (p. 100), by National Research Council, National Committee on Science Education Standards and Assessment, 1996, Washington, DC: National Academy Press.

fundamental science concepts. Students will gain a deeper understanding and appreciation of science if they cover fewer topics and instead "uncover" some in greater depth (i.e., "less is more"; Labov, 2006; NRC, 1996). The following section outlines the observable practices of effective elementary science instruction that will be examined in this study.

Observable and Measurable Reformed Teaching Practice

In a synthesis of findings report, *Taking Science to School* (2007), the NRC emphasized five key teaching models and practices observable and measurable in effective elementary science instruction: (a) talk and argument, (b) modeling and representations, (c) investigations and inquiry, (d) alignment to science core concepts, and (e) appropriately addressing science misconceptions. These strands rely on constructivist principles that, in the context of science instruction, are observable by evaluators and measurable through the use of classroom observation instruments. The following sections explain these constructs and discuss the role they play in effective elementary science instruction.

Talk and Argument

The NRC (2007) suggested that effective science instruction includes making students' thinking visible through a construct called "talk and argument." In the past two decades, the role of language in the science curriculum has become prominent in science education literature (Dawes, 2004; Gee, 1989; Lemke, 1990; Yore, Bisanz, & Hand, 2003). From a constructivist perspective, language mediates social interaction and meaning is constructed as learners interpret and reinterpret events through the lens of prior knowledge (Barnes, 1992; Berk & Winsler, 1995; Vygotsky, 1986). In order to process, make sense of, and learn from their ideas, observations, and experiences, students must talk about their ideas. Talking is integral to science learning (Michaels et al., 2008).

Argument can be classified as rhetorical, dialectical, or analytical discourse (Duschl & Osborne, 2002). Rhetorical arguments are one-sided arguments used to persuade others by presenting one point of view as more convincing than the alternatives (Driver, Newton, & Osborne, 2000; Yore et al., 2003). Dialectical arguments, sometimes referred to as dialogical or multifaceted arguments, involve the examination of differing perspectives during discussion or debate. Analytical arguments follow the rules of logic (e.g., Toulmin, 1958) and may be inductive or deductive (Duschl & Osborne, 2002; Yore et al., 2003). Inductive arguments include analogies and causal correlations, while deductive arguments include syllogisms and causal generalizations (Duschl & Osborne, 2002). Current science education reform emphasizes the use of dialectical and analytical arguments while deemphasizing rhetorical arguments, which traditionally have been predominant in the classroom (Driver et al., 2000).

Effective instruction in talk and argument should include a high ratio of interactions between students and the teacher (NRC, 2007). For example, when posing a science question, the teacher should actively extend student thinking or ask the students to support their claims with evidence. Effective instruction should also include frequent opportunity for students to discuss scientific ideas with their peers (NRC, 2007). While facilitating talk and argument in the classroom, effective teachers should also provide relevant examples and analogies and demonstrate accurate science language in their teaching (Michaels et al., 2008; NRC, 2007; Tippett, 2009). These teaching practices would be observable and measurable in effective science instruction.

Modeling and Representations

In the past two decades, science education experts have increasingly recognized the value of modeling and representations in the science education reform movement (AAAS, 1993; Giere, 1991; Gobert & Buckley, 2000; NRC, 1996; NRC, 2007). At present, models and modeling are considered integral parts of scientific literacy (S. Gilbert, 1991; J. Gilbert, 1993; Gilbert & Boulter, 1998; Linn & Muilenberg, 1996; Perkins, 1986). The general definition of models put forth by Ingham and Gilbert (1991) is that a model is a simplified representation of a system, which concentrates attention on specific aspects of the system. Furthermore, models enable complex and abstract aspects of the system to be rendered either visible or more readily visible (J. Gilbert, 1995). For example, a teacher could create an atom model with a solar system like display, with the nucleus surrounded by electrons. This would be a simplified representation of the atomic system; however, it would make the abstract aspects of an atom visible to students.

Scientists develop models and representations as ways to think about the natural world. The kinds of models and representations scientists use vary widely. In a general sense, a model is a representation of a phenomenon, an object, or idea (Gilbert et al., 2000). In science, a model is the outcome of representing an object, phenomenon or idea (the target) with a more familiar one (the source; Tregidgo & Ratcliffe, 2000). The model can only relate to some properties of the target; some aspects of the target must be excluded from the model (Driel & Verloop, 1999). A model may be a prototype for a whole class of similar things. For example, the solar system model of the atom displays the nucleus surrounded by electrons but excludes the delocalization of electrons. This model will represent some properties of an atom; however, many aspects of a specific kind of atom cannot be included in the model.

There are different types of models in science education. To categorize them, one should understand the difference between conceptual and mental models. Conceptual models are devised as tools for the understanding or teaching of systems (Ornek, 2008). In addition to this, conceptual models are external representations—socially constructed and shared—that are precise, complete, and consistent with the shared scientific knowledge specially created to facilitate the comprehension or the teaching of the systems in the world (Greca & Moreire, 2000). Conceptual models include mathematical models, computer models, and physical models (Ornek, 2008).

On the other hand, mental models are what people have in their heads that guide their use of things (Norman, 1983). Vosniadou (1994) defined mental models as analog representations that preserve the structure of the thing they represent. Buckley and colleagues (2004) also viewed mental models as internal, cognitive representations. Mental models have a variety of features: (a) mental models are generative, (b) mental models involve tacit knowledge, (d) mental models are synthetic, and (d) mental models are restricted by worldviews (Franco & Colinvaux, 2000).

First, mental models are generative, which means that people or students can produce new information and make predictions while they are using mental models (Franco & Colinvaux, 2000). Next, mental models involve tacit knowledge; a person using a mental model is not completely aware of some aspects of his or her mental models (Franco & Colinvaux, 2000). Also, mental models are synthetic, or are simplified representations of the target system (Franco & Colinvaux, 2000). Finally, mental models are constrained by worldviews, meaning that people develop and use mental models according to their beliefs (Franco & Colinvaux, 2000). In other words, a set of limitations constrains the possible mental models that people use.

Representation is a predecessor to full-fledged modeling. Even very young children can use one object to stand in for or represent another (NRC, 2007). However, young children typically do not recognize or account for the relationships and separations between the real world and models. Also, a child may have difficulty differentiating the features of a phenomenon that a representation accounts for or fails to account for (Michaels et al., 2008). The use of all forms of symbolic representation, such as graphs, tables, mathematical expressions, and diagrams, can be developed in young children and lead to more sophisticated modeling in later years.

Representations or symbolic development can serve as an important guide for incorporating modeling into instruction. Lehrer and Schauble (2004) observed characteristic successions in the understanding of modeling over the span of the elementary grades. They developed a learning progression that emphasizes different and increasingly complex ideas for different ages of children. Teachers must make informed decisions about how and when to introduce increasingly challenging forms of models to support science learning over the long term through using models and representations (Lehrer & Schauble, 2006).

Effective science instruction includes making students' thinking visible through modeling and representations (NRC, 2007). Effective teaching indicators within the modeling and representation construct include: (a) the teacher using models to demonstrate concepts, (b) the teacher using models to assess student understanding, and (c) the teacher having students demonstrate their understanding of science through science writing, drawings, and mathematical data representations (Michaels et al., 2008).

Investigations and Inquiry

Recent studies suggest that effective science instruction includes learning science through investigations and inquiry-based instruction (NRC, 2007). But what exactly is inquiry-based instruction? Cuevas, Lee, Hart, and Deaktor (2005) argued that this question is difficult to address because there is no clear or agreed-upon conception of what science inquiry involves. Researchers who study inquiry-based instruction are not in favor of a predetermined procedure for teaching; however, they do agree that a systematic study of inquiry methods is necessary in promoting its use in science education reform (Crawford, 2000; Wu & Krajcik, 2006). Other researchers noted that although the literature did attempt to define inquiry, it did little to prescribe how to conduct inquiry in a classroom (Crawford, 2000; Keys & Bryan, 2000; Wu & Krajcik, 2006).

Inquiry teaching is difficult to characterize because the method has several modes and levels. Inquiry-based instruction varies in form (open vs. closed), in its locus of control (teacher-centered vs. student-centered), and in its magnitude (simple vs. complex); however, its function is constant (Richardson & Liang, 2008). The function of inquiry-based instruction is for students to find answers to questions by way of gathering data or evidence. Inquiry is a process through which scientists attempt to find answers to questions through observation, exploration, experimentation, and investigation.

Even though researchers do not advocate a prescriptive approach to inquiry-based instruction, the construct must be defined before it can be systematically examined (Richardson & Liang, 2008). The NRC provides some general guidelines for inquiry-based instructional practices that elementary teachers should employ during classroom investigations. The NRC (2007) indicators of effective teaching using science inquiry and investigations include: (a) science investigations are directed by the teacher in small groups, (b) science investigations are student centered, and (c) core science concepts are taught and assessed within the investigation.

Student-centered, small-group science investigations provide an engaging way for students to develop a strong grasp of science content, the practices of scientific work, and the nature of science itself (Michaels et al., 2008). Investigations should utilize the four strands of science learning as outlined in the NRC report, *Taking Science to School* (2007). These four strands include: (a) understanding science explanations, (b) generating scientific evidence, (c) reflecting on scientific knowledge, and (d) participating productively in science (NRC, 2007).

At the root of all science investigations are complex problems and compelling questions. In order for problems and questions to be effective for supporting science learning, they must be meaningful from the perspective of the student as well as from the perspective of science as a discipline (Michaels et al., 2008). Effective science investigations should be aligned with grade level core concepts, and understanding of these concepts should be assessed within the investigations. For example, a fourth grade teacher could design an investigation of water conservation in which students would gather data about the amount of water needed for different types of plants and decide which plants would survive best in a desert environment.

Supporting student learning in regard to scientific investigations requires deliberate and consistent instructional effort. Research shows that simply "doing" science activities often leaves students with an inaccurate idea of what science is and how science works (Michaels et al., 2008). In contrast, effective investigations employ the four strands of science learning by requiring student to gain understanding of science explanations, generate scientific evidence, and then reflect on their scientific knowledge. These investigations provide opportunities for students to participate productively in science by working in small groups and presenting their findings to peers. Building students' knowledge and skills across the strands requires intentional, sustained instruction and support.

Investigations that support student learning require teachers who understand how specific problems evolve, and teachers themselves will need to have first-hand experiences akin to those they create for their students (NRC, 2007). Schools, universities, foundations, science centers, museums, government agencies, and professional development programs must find ways for teachers to have these experiences, building their knowledge and comfort level with science in order to create an effective environment for student learning (Michaels et al., 2008).

Science Core Alignment

In the standards-based reform movement, science core curriculum drives classroom instruction. The Utah Elementary Science Core Curriculum was designed using the American Association for the Advancement of Science's Project 2061: Benchmarks For Science Literacy (AAAS, 1993) and the National Academy of Science's NSES (1996) as guides to determine appropriate content and skills. The Utah Elementary Science Core described "what students should know and be able to do at the end of each grade level. It was developed, critiqued, piloted, and revised by a community of Utah science teachers, university science educators, State Office of Education specialists, scientists, expert national consultants, and an advisory committee representing a wide variety of people from the community" (Utah State Office of Education, 2002, ¶2).

The Utah Elementary Science Core reflects the current philosophy of science education expressed in national documents and has the endorsement of the Utah Science Teachers Association. The Core reflects high standards of achievement in science for all students. The NRC (2007) recommended that effective science instruction align with science content. The indicators of effective instruction within the science content alignment construct require that the teacher identifies and clearly stated the learning objective and aligns instruction with the objective.

Addressing Science Misconceptions

Effective science instruction requires the teacher to appropriately address science misconceptions when they occur. One of the great pleasures of working with children is their enthusiasm and lack of inhibition in creating and considering new ideas. Recent research has revolutionized views of how children's minds develop from infancy through adolescence. The past 20 to 30 years of research have shown that children come to school with a great capacity for learning in general and learning science in particular (Metz, 1995; NRC, 1999, 2007). Children typically have significant gaps in their understanding (as do many adults), and their unschooled reasoning abilities may lead them to draw erroneous conclusions (Michaels et al., 2008). But children are not the bundles of misconceptions they are sometimes portrayed as being. They are active explorers who have successfully learned about regularities in particular domains of experience in ways that help them interpret, anticipate, and explain their world (Metz, 1995; NRC, 1999, 2007).

Science education is sometimes seen as a process of filling students up with facts.

According to this line of reasoning, if students learn enough concepts, definitions, and discrete facts, they will understand science. Learning new facts is important in science education, but learning facts alone is not enough. To understand science, children also need to view facts in broader contexts of meaning (Michaels et al., 2008). They need to reposition their prior knowledge within a larger network of ideas and learn how to think about scientific explanations (NRC, 2007).

When learning complex material, such as concepts encountered in a science lesson, a student can experience at least three different conditions of prior knowledge. Chi (2008) described these three conditions as follows. First, a student may have some related knowledge but no prior knowledge of the new concepts. In this case, prior knowledge is missing, and learning consists of adding new knowledge. Second, a student may have some correct prior knowledge about the to-be-learned concepts, but that knowledge is incomplete. In a third condition, a student may have prior knowledge, either from school or everyday experience, which is in conflict with the to-be-learned concepts. Knowledge acquisition under this third case is of the conceptual change kind (Vosniadou, 2004). "Thus, learning in this third condition is not adding new knowledge or gap filling incomplete knowledge; rather, learning is changing prior misconceived knowledge to correct knowledge" (Chi, 2008, p. 61). Researchers group these kinds of changes in thinking into the general category of conceptual change (Chi, 2008; diSessa & Minstrell, 1998; Vosniadou, 2004).

The elementary and middle school years can include impressive periods of conceptual change. Children can have dramatic new insights that change the way they understand a whole domain, developing new understandings that change their lives (Michaels et al., 2008). Conceptual change of the kind that is needed in K-8 science instruction can be difficult to engineer. Many teachers have their students do experiments or make observations with the hope that scientific understanding will miraculously emerge from the data (Metz, 1995). Being exposed to new information, however, is not the same as understanding or integrating that information into what one already knows. Real conceptual change requires that deeper reorganizations of knowledge occur (diSessa & Minstrell, 1998). In order for teachers to appropriate address science misconceptions they need to facilitate real conceptual change through discussion of how to understand the concept. To address science misconceptions, the teacher must clarify the misconception and provide the students with insight that allows for conceptual change and correction of the misconception (Michaels et al., 2008).

Effective elementary science instruction requires much from a teacher. Reformed science education calls for instruction that aligns to core concepts and appropriately addresses science misconceptions; it also includes talk and argument, modeling, and investigation. Effective science instruction takes time to develop. Well-designed professional development programs can support reformed teaching practice by preparing teachers for these reformed teaching models and practices. The following section outlines the professional development characteristics and components that have proven effective in reforming the classroom science instruction of elementary school teachers.

Professional Development in Elementary Science Education

Well-designed opportunities for teacher learning can produce desired changes in classroom practices, can enhance teacher's capacity for continued learning and professional growth, and can in turn contribute to improvements in student learning (American Federation of Teachers, 2002; NRC, 2007; National Staff Development Council [NSDC], 2001). In general, a great deal is known about the characteristics of effective professional development. There is an overall consensus about these characteristics among researchers and among professional and reform organizations (AFT, 2002; Elmore, 2002; Knapp, McCaffrey, & Swanson, 2003; NRC, 2007; NSDC, 2001).

Research has identified features of quality teacher learning opportunities that can be realized through a wide variety of organizational structures including mentoring and coaching, teacher work groups, and expert-led programs of professional development (Appleton, 2003; Johnson, Kahle, & Fargo, 2007; Loucks-Horsley, 1998; NRC, 2007; Sparks & Hirsch, 1997). Drawing heavily on three previous attempts to synthesize the literature on effective teacher development (American Educational Research Association, 2005; Elmore, 2002; Odden, Archibald, Fermanich, & Gallagher, 2002), the NRC has created a list of seven significant features. Research suggests that well-structured opportunities for teacher learning:

1. Reflect a clear focus on the improvement of student learning in a specific content area that is grounded in the curriculum they teach.
2. Focus on the strengths and needs of learners in the setting and evidence about what works drawn from research and clinical experience.

3. Include school-based and job-embedded support in which teachers may engage in assessing student work, designing or refining units of study, or observing and reflecting on colleagues' lessons.

4. Provide adequate time during the school day and throughout the year, including considerations of the time required for both intensive work and regular reflection on practice. Furthermore, the overall span of time for teacher professional development is several years.

5. Emphasize the collective participation of groups of teachers, including opportunities for teachers from the same school, department, or grade level.

6. Provide teachers with a coherent view of the instructional system (e.g., helping teachers see connections among content and performance standards, instructional materials, local and state assessments, school and district goals, and the development of a professional community).

7. Require the active support of school and district leaders. School leaders who participate in creating and sustaining teacher learning opportunities are better positioned to support teachers' use of new knowledge and skills. (NRC, 2007, p. 307)

These features provide a frame for describing, comparing, and analyzing the organization of teacher learning across schools, districts, and within preservice and inservice teacher development programs.

Among the more rigorous studies of professional development for teachers of science are those of a longitudinal study of sustained professional development by the Merck Institute for Science Education (Corcoran, McVay, & Riordan, 2003); the NSF-funded studies of systemic reform in mathematics and science (Supovitz & Turner, 2000; Weiss et al., 2003); and evaluations of the federal Eisenhower mathematics and science professional development program (Garet et al., 1999). These studies have examined both elementary and secondary science education. This study focused on research in elementary science education.

Elementary science education differs from secondary science education. Effective professional development programs in elementary science education must acknowledge and address these differences. Elementary science is integrated, covering concepts from many different areas of science. Current K-8 science standards include objectives in earth and space science, life science, physical science, technology, practices of science, nature of science, and history of science (NRC, 1996). Whereas secondary science teachers can focus on one area of science, elementary teachers must incorporate multiple areas. Therefore, professional development for elementary teachers should address the complexity of elementary science instruction.

The growing importance of science in the modern world has focused increasing attention on elementary science education. The development of the national standards and benchmarks in the 1990s catalyzed a nationwide conversation about what students need to learn in science and how elementary science education can support that learning (AAAS, 1993; NRC, 2007; NSES, 1996; NSTA, 2003). Recently, professional development programs have been designed specifically for elementary teachers. These programs aim to incorporate the initiatives of elementary science education reform.

Elementary school teachers, like other learners, need support. Teachers and science experts on the district, state, and university levels must build partnerships to support the exchange of knowledge and information related to core concepts of science across grade levels (NRC, 2007). Effective professional development programs honor elementary teachers while providing support for them through partnerships with outside science experts. In achieving this balance, Carlone and Webb (2006) contended, "Innovative approaches to professional development take seriously teachers' knowledge, goals, context, voice, and experience" (p. 546). These approaches do not give absolute authority to the teacher, nor do they completely accept the authority of the outside expert (Campbell, 2012; McIntyre & Hagger, 1992). These approaches build a network of science educators and professional development experts to work together to ensure that the complex instructional practices required for science education reform are supported by systematic, sustained professional learning throughout a teacher's career (Duschl et al., 2006; Michaels et al., 2008).

New frameworks and standards documents point toward a kind of elementary science instruction that differs substantially from what occurs in most classrooms today (Michaels et al., 2008; NRC, 2007; NSES, 1996). The new vision of elementary science education embraces different ways of thinking about science, different ways of thinking about students, and different ways of thinking about elementary science education (Michaels et al., 2008). This vision of science education is best accomplished by viewing the professional development of teachers as a transformational process (Campbell, 2012; Wenger, 1998). Several studies list common qualities of effective professional development including an emphasis on content knowledge, collaboration with experts and peers, time for reflection, and sustained teacher learning over time (Davis, 2003; Supovitz & Turner, 2000; van Driel et al., 2001). These qualities align with transformational learning concepts and provide a way for teachers to reform their science learning and teaching practice.

NSES outlined the new vision of elementary science education through promoting changing emphases in professional development. The NSES initiatives and standards of preservice and inservice professional development will be the focus of the next section.

National Science Education Standards for Professional Development

The NSES call for fundamental changes in what teachers should know and be able to do, especially for elementary and middle school teachers, who increasingly are becoming teachers of science (Labov, 2006; NRC, 1996). These recommendations suggest that teachers need new and very different approaches to teacher preservice preparation and ongoing inservice professional development. Table 2.3 contains excerpts from NSES of standards for the professional development of teachers of science. Table 2.4 contains the changing emphases from NSES in professional development.

As illustrated in these tables, the NSES' changing emphasis in science teaching and standards for professional development suggest the need for professional

Table 2.3

Excerpts from NSES of Standards for the Professional Development of Teachers of Science

Standard	Excerpt
A	The professional development of teachers of science requires learning essential science content through the perspectives and methods of inquiry.
B	Professional development of teachers of science requires integrating knowledge of science, learning, pedagogy and students; it also requires applying that knowledge to science teaching.
C	The professional development of teachers of science requires building understanding and ability for lifelong learning.
D	Preservice and inservice professional development programs for teachers of science must be coherent and integrated.

Note. From *National Science Education Standards* (p. 59-68), by National Research Council, National Committee on Science Education Standards and Assessment, 1996, Washington, DC: National Academy Press.

Table 2.4

The NSES Changing Emphasis on Professional Development

Less emphasis on…	More emphasis on…
Transmission of teaching knowledge and skills by lectures	Inquiry into teaching and learning
Learning science by lecture and reading	Learning science through investigation and inquiry
Separation of science and teaching knowledge	Integration of science and teaching knowledge
Separation of theory and practice	Integration of theory and practice in school settings
Individual learning	Collegial and collaborative learning
Fragmented, one-shot sessions	Long-term coherent plans
Courses and workshops	A variety of professional development activities
Reliance on external expertise	Mix of internal and external expertise
Staff developers as educators	Staff developers as facilitators, consultants, and planners
Teacher as technician	Teacher as intellectual, reflective practitioner
Teacher as consumer of knowledge about teaching	Teacher as producer of knowledge about teaching
Teacher as follower	Teacher as leader
Teacher as an individual based in a classroom	Teacher as a member of a collegial professional community
Teacher as target of change	Teacher as source and facilitator of change

Note. From *National Science Education Standards* (p. 72), by National Research Council, National Committee on Science Education Standards and Assessment, 1996, Washington, DC: National Academy Press.

development to be long-term, focused on building teacher science literacy, rooted in research on how children learn science, and embedded in teaching practice in the classroom (NRC, 1996). These characteristics will be explored further in the following sections.

Sustained Professional Development

Adult transformational learning takes time (Mezirow, 2000). Peacock and

Rawson (2001) suggested that sustainable professional development and reform occur when teachers are able to identify their own set of competencies for effective teaching and have time to work toward self-defined goals. Sustained professional development can be a way for teachers to become aware of their learning and can provide a transformational learning experience, through which they can assess and improve their own methods, knowledge, and skills (Loucks-Horsey et al., 2003).

Science education reform depends upon engaging teachers in sustained professional development (NRC, 1996) and developing teachers who view themselves as life-long learners. Teacher learning should parallel the concepts of constructive student learning by continually building and reflecting on what has been learned (Davis, 2003; Osborne, 1998). For this reason the NRC recommends that professional development for elementary science education be long-term, sustained for 3 years or more (NRC, 2007).

How Children Learn Science

In the book, *How students learn: Science in the classroom* (NRC, 2005), three principles of science learning are described. The three principles include: (a) addressing preconceptions, (b) knowledge of what it means to “do science,” and (c) metacognition. These three principles have guided the science education reform efforts of the NRC. Each principle will be described briefly below.

The NRC’s first principle emphasizes the importance of addressing preconceptions in how children learn science. Students bring conceptions of everyday phenomena to the classroom that are quite sensible, but scientifically limited or incorrect. Teachers must address those ideas through classroom argumentation and reasoning if students are to understand science. Understanding scientific knowledge often requires a change in what people notice and understand about everyday phenomena. Effective science instruction provides students with many rich opportunities to experience and understand phenomena from new perspectives (Carey, 2000; NRC, 1999). These experiences encourage students to change their noticing, thinking and understanding while learning science.

The NRC's second principle emphasizes the importance of knowing what it means to do science. Science teaching reform efforts rely heavily on constructivist principles. At the heart of constructivist philosophy is the belief that knowledge is not given but gained through meaningful experiences (Piaget & Inhelder, 1969; Vygotsky, 1978). In science, these meaningful experiences often come through authentic inquiry (NRC, 2005). Authentic inquiry involves observation, imagination, and reasoning about the phenomena under study. It includes the use of tools and procedures that encourage students to extend their everyday experiences of the world and help them organize data in ways that provide new insights into phenomena (Petrosino, Lehrer, & Schauble, 2003).

The NRC's third principle of how students learn science emphasizes metacognition. Metacognition is the awareness or analysis of one's own learning or thinking processes (Briggs, 1987). Much of the research on metacognition focuses on comprehension of text. Comprehension of text clearly applies to science, where text can be complex and difficult for many students to comprehend. However, metacognition is not simply comprehension, it includes monitoring of and reflection on scientific reasoning (NRC, 2005). According to the NRC, "being metacognitive about science is different from simply asking whether students comprehend what they read or hear; it requires taking up the particular critical lens through which scientist view the world" (p. 410).

Science Content

Professional development should align with the science content that teachers teach and include opportunities to learn about science concepts as well as common misconceptions. The professional development literature asserts that teacher learning should parallel the experiences that reformers want students to receive from these teachers (Loucks-Horsley et al., 2003; van Driel et al., 2001). Therefore, teachers need to be very familiar with the science concepts that should be taught in their classroom. They also need to be aware of commonly held misconceptions about these concepts.

Many elementary teachers, like many college-educated professionals, have only a superficial knowledge of science. Inadequate undergraduate coursework and insufficient professional development opportunities contribute to the problem (NRC, 2007). Mounting evidence suggests that what a teacher knows about science influences the quality of instruction and has a powerful effect on the success and quality of instruction that teachers can provide for students (Loucks-Horsley et al., 2003; van Driel et al., 2001). A teacher's continuous construction of subject matter knowledge determines, in part, what happens in the classroom. A study by van Driel and colleagues (2001) reported that the teachers' level of science subject matter knowledge correlated directly with the number of science teaching strategies employed by the teacher.

In order to teach science effectively, the teacher must first understand the subject matter being taught (Michaels et al., 2008). Teachers must demonstrate science content literacy and deep understanding of core science concepts in order to teach science effectively (NRC, 2007). Effective professional development should continually increase the science literacy of participating teachers.

One way that teachers can be aware of their own learning and incorporate it into their teaching is through participating in professional development. Several studies list common qualities of effective professional development as having an emphasis on content knowledge, collaboration with experts and peers, time for reflection, and sustained teacher learning over time (Davis, 2003; Supovitz & Turner, 2000; van Driel et al., 2001). These qualities align effective professional development with the transformational learning theory and provide a way for teachers to reform their science learning and teaching practice.

Teaching Practice

Professional development is most effective when teachers develop reformed practices through a process shaped by standards and knowledge gained from classroom practice (Birman et al., 2000; Lemke, 2001; NRC, 1999, 2011; Stein et al., 1999; Wenger, 1998). Innovative approaches to professional development encourage teachers to examine basic questions about what it means to be a teacher through connecting the professional development experience with instructional practice in the classroom (Campbell, 2012; NRC, 2003). Current research in teacher development supports the extensive integration of professional development experiences with classroom instruction (Birman et al., 2000; Lemke, 2001; NRC, 1999, 2011; Stein et al., 1999). When this integration is effective, it allows participating teachers the opportunity to examine the basis of what is being learned in the context that it will be employed (Campbell, 2012).

The NSES changing emphasis in science education suggests the need for professional development to be embedded in teaching practice in the classroom. Professional development should extend to the classroom through partnerships and professional learning communities (NRC, 2007). It should model the reformed practices that teachers are learning to implement in their classrooms (Birman et al., 2000; Lemke, 2001; NRC, 1999, 2003; Stein et al., 1999; Wenger, 1998).

Science education reform calls for specific changes called "best practices" to be utilized in classroom instruction. Loucks-Horsley and colleagues (2003) described characteristics of best practices in science education reform by using five goals.

- The tasks presented by the teachers are familiar ones to the students, allowing them to mobilize their prior knowledge and tap into what they already know about the new concept that they are learning.

- Students demonstrate an extraordinary level of engagement. The students will be able to explain what they are doing and why they are doing it.

- The tasks in which the students are engaged are constructed around significant concepts in science. The tasks require the students to think scientifically by making conjectures and hypotheses and marshaling different forms of evidence to support or refute them.

- Students are communicating constantly with each other. Students are not restricted to consider only the ideas of the teacher and the textbook authors. Students' understanding of science deepens and develops through communication and community.

- Teachers act as coaches establishing a classroom environment that enables students to express themselves. Teachers monitor classroom interactions, deciding when and how to intervene. Teachers help students make sense of their observations and analyze carefully what is going on. (p. 6)

When trying to transform classroom instruction, these five goals provide a framework for professional development in elementary science education. Effective programs will heed the call for professional development to be embedded in teaching practice and promote best teaching practice in the classroom (NRC, 1996).

To determine the effectiveness of a particular professional development program, participating teachers must be observed and rated by trained evaluators using an instrument designed to determine the degree to which classroom instruction matches the instructional models provided in the professional development. The professional development program under investigation for this study is based on recommendations from the NRC. The next section summarizes three recent studies that examined the effects of elementary science professional development on classroom practice.

Recent Studies

Recent studies on the effects of teacher development on classroom instruction report mixed findings on the effectiveness of teacher development on changing elementary science teaching practice. The following studies reveal the current state of understanding about elementary science teacher professional development. The first study by Minuskin (2009) researched the effects of professional development on the knowledge and classroom practices of teachers of science in fourth grade. The researcher implemented an 18-week professional development program that used a collaborative model involving eight teachers who were all from the same school district in New Jersey. Before and after the intervention, the researchers observed instruction and used Horizon Research's observation protocol (Horizon Research, Inc., 2001). This instrument is used to measure science teaching effectiveness pre and post intervention and has a combination of checklist and 5-point scales.

Findings from this study suggest that teacher instruction did not significantly differ after the professional development intervention. Minuskin's study (2009) also suggests that teacher content knowledge did not significantly increase due to the intervention. This study was limited in length of treatment and number of participants. Minuskin speculated that local factors influenced the outcome. Local factors included poor teacher stability and morale as well as changes in leadership and programs. In the year of intervention there was a search for a new superintendent, 64 tenured teachers were dismissed from their job, and three new reading and math initiatives were introduced in the district. Minuskin recommended that future studies include a more systemic program that involves a more diverse group of teachers and a wider range of stakeholders from across a state or region; thereby reducing the influence of local factors.

In the second study, Santau (2008) conducted a 1-year research project that examined teachers' knowledge and practices in science instruction with English language learning (ELL) students. The study participants were 32 third-grade, 21 fourth-grade, and 17 fifth-grade teachers in the first-year implementation of the intervention. Classroom observations were conducted using observation protocols adapted from Horizon Research's observation protocol instrument (Horizon Research, Inc., 2001). Results indicated that teachers' knowledge and practices were within the bounds of the intervention, but short of reform-oriented practices, and that relationships among the four domains existed, especially for teachers in grade five. The four domains included: (1)

teachers' knowledge of science content, (2) teaching practices to promote scientific understanding, (3) teaching practices to promote scientific inquiry, and (4) teaching practices to support English language development during science instruction. The study reported findings from only the first-year implementation of a multi-year professional development intervention. Santau (2008) suggested future research include the continuation of the intervention and its impact on change with teachers over three or more years.

A third study by Drits and Stark (2011) explored teacher change during a 1-year reform-based professional development program. The researcher also explored changes in teaching the year following treatment. The study examined patterns of change in elementary teachers' inquiry practices, inquiry beliefs, and physical science content knowledge during both years, as well as the effects of school-level and individual-level factors on these changes in the year following the program. Fifteen fourth- through sixth-grade teachers from three low-performing elementary schools participated in the study.

To measure whether or not teachers engaged students in inquiry and to what degree, Drits and Stark (2011) used the Reform Teaching Observational Protocol (RTOP) (Piburn et al., 2000). The findings indicated that the program was effective in advancing teacher change during the program year; scores in all three measures (inquiry practices, inquiry beliefs, and physical science content knowledge) increased at statistically significant rates. While scores increased in all three measures during the year following the professional development program, only content knowledge scores increased significantly. The study population was small and consisted of all volunteers, and only three schools participated in this study. Because the number of teachers and schools studied was small, it is difficult to generalize from the conclusions.

These three studies reveal the mixed findings and current understanding of the effectiveness of teacher development on changing elementary science teaching practice. The limitations of these studies included limited length of treatment, small number of participants, and little variety of schools. None of these studies included a control group. Researchers recommended that future studies include a long-term systemic program that involves a more diverse group of teachers from different schools across a state or region. The current study moves this research forward by addressing these limitations and recommendations through examining a comprehensive 3-year professional development program that involves teachers from 10 treatment schools from four districts across the state of Utah. The study also includes three control schools from three districts across the state of Utah. Twenty-two teachers, randomly selected from a treatment population of 148 teachers, were measured over time. The treatment sample was also compared to a control group of 20 nonparticipants.

Building on the foundation of prior research and published literature, this study examined the effects of a specific long-term professional development program on the science instruction of participating teachers. The transformational learning theory served as an appropriate theoretical lens because it is oriented in explaining the way adult learners interpret and reinterpret their views and practices (Mezirow, 1991). Its social constructivist orientation relates to professional development wherein groups of teachers and experts construct knowledge collaboratively by creating a culture of shared practices and meanings (Mezirow & Taylor, 2009). The transformational learning theory (Mezirow, 1991) frames the professional development program under examination as well as the research questions and methodology of this study.

In order to meet the objectives of this study, the researcher prepared by gaining a full understanding of the professional development program through participation and study over 2 years. A brief description of the professional development program under examination seems appropriate at this point. Specific attention will be given to how the theoretical lens ground this professional development model and how it is supported by research and standards aligned reform.

Partnership for Effective Science Teaching and Learning

The Partnership for Effective Science Teaching and Learning (PESTL) is a 3-year professional development program funded by a grant from the Utah State Office of Education under the U.S. Department of Education's Elementary and Secondary Education Act (ESEA) Title II part B Mathematics and Science Partnership grant program. The overarching objective of PESTL is to improve student learning through sustained teacher professional growth and science literacy. PESTL is based on recommendations from the NRC's report *Taking Science to School: Learning and Teaching Science in Grades K-8* (2007). It is designed to help the elementary practitioner improve science teaching and learning through utilizing partnerships and professional learning communities (PLCs).

PESTL is designed to focus on teachers as adult professionals, concentrating on teacher needs and professional ways of thinking and learning. The professional development is based on transformational learning theory concepts and includes activities that are learner-centered, useful, and collaborative. The workshops focus on the significance of science in everyday life and the importance of teaching science to elementary age children.

PESTL utilizes public partnerships with the Utah State Office of Education and four school districts across the state of Utah. The program also maintains partnerships with two universities: Weber State University and Southern Utah University. As a business partner, Merck Institute for Science Education (MISE) donates copies of *Ready, Set, SCIENCE!* to all PESTL participants.

PESTL has five professional development components that work together to sustain the effectiveness of the program over time. They are:

- Summer Seminars
- School Professional Learning Communities (PLCs)
- Grade-Level Alignments
- Mid-Winter Institutes
- Content Courses

A summary of the implementation schedule for each year of the program is illustrated in Table 2.5. The following descriptions will give the reader an overview of each component.

Table 2.5

Summary of PESTL Activities

Component	Partnerships	Year 1	Year 2	Year 3
Summer seminars	District University MISE	5 days in summer	5 days in summer	5 days in summer
School PLCs	School MISE	8 times during school year	6 times during school year	4 times during school year
Grade level alignments	District University	2 afterschool sessions during school year	2 afterschool sessions during school year	3 afterschool sessions during school year
Mid-winter institutes	District University	1 day each year	1 day each year	1 day each year
Content courses	District University	1 day each year in the fall	1 day each year in the fall	1 day each year in the fall

Summer Seminars

Summer seminars run for 5 days (Monday-Friday) and utilize partnerships with university science departments and other experts in science education to provide the week's schedule of professional development workshops. A major objective of the summer seminar is to further develop the teacher's science literacy and to utilize "conceptual models" for understanding science phenomena.

Over 3 years, these seminars focus on different aspects of teaching science as outlined in *Taking Science to School: Learning and Teaching Science in Grades K-8* (NRC, 2007) and the complementary practitioners book, *Ready, Set, SCIENCE!* (Michaels et al., 2008). The first year the seminar focuses on making students' thinking visible through talk and argument. The second year the seminar focuses on learning from science investigations. The third year the seminar focuses on making students' thinking visible through useful models and representations. For the summer seminars, presenters utilize the four strands of science learning (Michaels et al., 2008) and model best teaching practiced through process-oriented lessons.

School PLCs

The PESTL PLCs are held at individual schools and are facilitated by a fellow teacher. The facilitating teacher is selected through recommendations from the school principal and observations PESTL professional developers make during the first summer seminar. Over the 3-year cycle of the program, PLC meetings are held eight times the first year, six times the second year, and four times the third year. The PLC activities include presession readings, open discussions, lesson ideas, and self-reflective journal writing. Most of the readings for the PLCs come from *Ready, Set, SCIENCE!* (Michaels et al., 2008). PLCs are structured meetings that follow a set outline for learning. The facilitator keeps records of the PLCs and communicates school successes and specific teacher needs to the PESTL program coordinators.

Grade-Level Alignments

The grade-level alignments are district-wide, grade-level specific sessions held to provide guidance and discussion on instructional activities, science concepts, and formative assessment. These sessions help teachers connect state science curriculum to current NRC recommendations to use core concepts over time to teach science (Michaels et al., 2008; NRC, 2007). Teachers meet two or three times each year after school to participate in these sessions. The teachers work together with a PESTL facilitator to align the grade level science standards with crosscutting concepts and models.

Mid-Winter Institutes

Mid-winter institutes focus on the nature of science and science practices. Teachers meet together district-wide each year for these one-day sessions. Mid-winter institutes include inquiry activities by grade level and, like the summer seminars, utilize partnerships with university science departments and other science experts to provide professional development workshops. Presenters also make use of the four strands of science learning (Michaels et al., 2008) and model best teaching practices while utilizing science investigations as learning activities.

Content Courses

The content courses are grade-level specific sessions that focus on the content of the Utah State Elementary Science Core Curriculum. Partnering university professors and other guest presenters teach these courses and facilitate them in a way that aligns core curriculum with reform-based science instruction, so that participants learn as they are asked to teach. The content course workshops are grounded in research on how children learn science and are presented in a classroom setting. Teachers from two districts meet together once a year for these 1-day sessions. The core content for each grade is organized into three sections, and the content courses focus on one section each year. For example, fourth-grade teachers are presented with content courses that focus on energy and the cycling of matter the first year; fossils, rocks and weathering the second year; and the interaction of living and nonliving things in an environment the third year.

Teachers also complete an online grade-level specific module on their own schedule. The online modules correspond with the science core curriculum covered in the content courses. The content courses and summer seminars offer university credit options through partnerships with local universities, and all PESTL sessions count toward teacher relicensure credits.

PESTL's five components support the need for professional development to be long-term, embedded in teaching practice in the classroom, and rooted in research on how children learn science as recommended by the NRC (1999). Each component is supported by research and standards aligned with elementary science education reform. Table 2.6 illustrates how PESTL aligns with current research and standards in science education.

Summary

This literature review established connections between professional development in elementary science education and the desired results of reformed classroom instruction. It established transformative learning theory as the theoretical lens for studying professional development in elementary science education. The review of literature identified professional development characteristics established significant in improving classroom science instruction of elementary school teachers. Additionally, five elements of effective instruction in elementary science education were identified and described. Three recent studies on the effects of teacher development on classroom instruction were explored, analyzed, and compared to this study. Finally, an overview of the PESTL professional development was given. This sustained professional development in elementary science education is the program under examination in this study. This researcher found that few quasi-experimental studies have investigated whether long-term teacher professional development in elementary science education actually yields changes in classroom instruction over time. Therefore, a study such as this is warranted.

Table 2.6

PESTL's Alignment to Science Education Reform Research and Standards Documents

PESTL component	Science education reform research and standards documents
Sustained teacher professional development	Several studies list a common quality of effective professional development as sustained teacher learning over time (Darling-Hammond & McLaughlin, 1995; Davis, 2003; Duschl et al., 2006; Garet et al., 2001; Loucks-Horsey et al., 2003; Michaels et al., 2008; NRC, 1996; NRC, 2007; Osborne, 1998; Peacock & Rawson, 2001; Sparks & Hirsh, 1997; Supovitz & Turner, 2000; van Driel et al., 2001).
Focus on scientific literacy	Several studies list a common quality of effective professional development as having an emphasis on content knowledge (Davis, 2003; Hodson, 2008; NRC, 1996, 2007; Supovitz & Turner, 2000; van Driel et al., 2001).
School PLCs	Several studies list a common quality of effective professional development as including collaboration peers and time for reflection (Davis, 2003; NRC, 1996, 2007; Supovitz & Turner, 2000; van Driel et al., 2001).
Embedded in teaching practice in the classroom Grade level alignments & content courses	Several studies list a common quality of effective professional development as having alignment with the science content that teachers teach (Davis, 2003; Loucks-Horsley et al., 2003; NRC, 2007; Supovitz & Turner, 2000; van Driel et al., 2001).
Utilization of partnerships in summer seminars & mid-winter institutes	Several studies list a common quality of effective professional development as including collaboration with science experts (Davis, 2003; NRC, 1996, 2007; Supovitz & Turner, 2000; van Driel et al., 2001).
Honoring teachers as adult professional learners	Several studies list a common quality of effective professional development as honoring and empowering teachers as adult professional learners (Campbell, 2012; Carlone & Webb, 2006; McIntyre & Hagger, 1992; NRC, 1996).
Rooted in research and standards documents on how children learn science	Several studies list a common quality of effective professional development as being grounded in research on how children learn science (Duschl et al., 2006; van Driel et al., 2001; Weiss et al., 2003; Yager, 2000).

CHAPTER 3

RESEARCH METHODS

Ms. King, now in her third year of professional development, feels she has made great progress in her science teaching and learning. A science expert has visited her class each year to observe her science instruction. She was nervous the first year, but now, with 2 years of professional development behind her, she feels confident and relaxed with the observer in her classroom. Ms. King has learned about using talk and argument, modeling, and investigations to make learning visible in her science lessons. She feels more effective in her overall science instruction. She is more confident in her own science literacy and feels more able to help her students when they have misconceptions about science concepts she is teaching them. Ms. King is delighted that she had the opportunity to participate in this sustained professional development program.

Introduction

This chapter discusses the methodology for this study. The main sections of this chapter are: study design, participants, measures, and procedures. The study design section describes the research design used for this study. The participants section describes the participants, sample, and control groups of this study and explains how these groups were selected. The measures section outlines the two classroom observation tools used in this study and describes how these instruments were created and validated. The procedures section describes the intervention and observations over the 3 years of treatment and gives an overview of the research questions and data analysis procedures for this study.

Study Design

The methodology for this study was quasi-experimental, using preexisting data from classroom observations. The researcher sought to determine the effects of sustained teacher professional development on the classroom instruction of elementary school teachers over a 3-year time period. This study sought to add to the current literature by examining the transformation of science instruction as the result of professional development designed to reform elementary science education. As mentioned before, few studies have examined changes in elementary science instruction over time as a result of sustained professional development.

To determine the overall effectiveness of the PESTL program, participating teachers were observed and rated by evaluators using classroom observation tools. Participants consisted of a cluster sampled treatment group of teachers who participated in all 3 years of the PESTL program. A control group of teachers, who did not participate in PESTL, was also evaluated for this study and served as a baseline for nontreatment. The data for the control group was collected during the first year of the study. PESTL evaluators conducted classroom observations in which 45 minutes of science instruction was examined during the regular school day. Two classroom observation instruments, the Summary Judgment of Science Instruction and the PESTL Observation Protocol on Instructional Effectiveness, were utilized to determine the degree to which classroom instruction was consistent with the instructional models provided in the professional development. These two instruments were used for both the treatment and control teacher observations.

The data used for this study were pre-existing, collected in classroom observations from 2008-2011. The researcher extrapolated a unique set of data from a larger set of data. The full data set included observation made throughout the 3 years of all 148 participating teachers. The data used for this study were selected based on cluster sampling of 22 teachers who participated in and were observed during all 3 years of the PESTL program. The full data set also included 53 control group teacher scores; however, none of these teachers were followed over time. Twenty control teacher observations were conducted during the first year. These 20 scores are the control condition and serve to illustrate a baseline for nontreatment. Hence, the unique set of data used by the researcher to conduct this study included 22 PESTL teacher scores from year one, year two, and year three, and 20 control group teacher scores from year one.

Participants

Participating teachers were cluster selected as the treatment sample for this study. In cluster sampling, instead of selecting all the subjects from the entire population, the researcher takes several steps in gathering the sample population (Lohr, 1999). In this study, 10 schools were selected using random sampling. Then from the 10 randomly selected schools, 22 teachers who participated in PESTL were randomly selected for observations. These 22 teachers were observed three times, once during year one of the treatment (2008-9), once during year two of the treatment (2009-10), and once in year three of the treatment (2010-11).

The participants of this study were 22 randomly sampled teachers who received PESTL professional development for 3 years. All participants were licensed elementary school teachers. The treatment sample teachers were from four districts across the state of

Utah. The four treatment districts were selected based on requirements set by the Utah State Office of Education in awarding the U.S. Department of Education's ESES Title II part B Mathematics and Science Partnership grant. The 16 treatment schools (four from each district) were selected based on student contextual factors. Seventy-five percent of the participating schools had high poverty and diversity characteristics, while the remaining 25% had average school composition. Table 3.1 outlines the selection of participating districts, schools, and teachers for this study.

The 22 participants of this study were fourth- through sixth-grade teachers. For the first year, 9 taught fourth grade, 10 taught fifth grade, and 3 taught sixth grade. In year two, one fourth grade teacher moved to fifth grade and one fifth grade teacher moved to fourth grade. For year three, the teacher who moved from fourth to fifth grade moved back to fourth grade. Five teachers were from district one, nine were from district two, seven were from district three, and one was from district four. This is a

Table 3.1

Selection of Participating Districts, Schools, and Teachers

State level	District level	School level	Teacher level
Four treatment districts were selected based on requirements set by the Utah State Office of Education in awarding the U.S. Department of Education's ESES Title II part B Mathematics and Science Partnership grant.	Sixteen treatment schools (four from each district) were selected based on student contextual factors. Three schools from each district were selected based on high need students and one school was selected based on average student composition.	All third through sixth grade teachers at each treatment school were invited to participate.	From the 16 schools, a total of 148 teachers participated in PESTL. Of these teachers, 22 participants were random cluster sampled for this study.

proportionate sample because district two had nearly twice as many participants as districts one and three, and district four had a very small number of participants. The number of participants in each district was affected by school size and teacher availability factors. Table 3.2 outlines the characteristics of the participating teachers.

Table 3.2

Characteristics of the Participating Teachers

Teacher number	District	Grade level year 1	Grade level year 2	Grade level year 3
1	1	4	4	4
2	1	4	4	4
3	1	5	5	5
4	1	5	5	5
5	1	5	5	5
6	2	4	4	4
7	2	5	5	5
8	2	6	6	6
9	2	4	4	4
10	2	5	5	5
11	2	5	5	5
12	2	6	6	6
13	2	4	4	4
14	2	4	4	4
15	3	4	5	4
16	3	4	4	4
17	3	6	6	6
18	3	4	4	4
19	3	5	5	5
20	3	5	5	5
21	3	5	5	5
22	4	5	4	4

The 20 control teachers were third- through sixth-grade teachers from the same four districts as the treatment teachers. One of the control teachers taught third grade, nine taught fourth grade, six taught fifth grade, and four taught sixth grade. Five teachers were from district one, six were from district two, eight were from district three, and there were no control teachers from district four. Table 3.3 outlines the grade levels and total numbers of participating and control teachers in this study.

The control group served to illustrate a baseline for treatment, because the treatment group was not observed prior to treatment. The control group consisted of a convenience sample of 20 third to sixth-grade teachers who did not participate in the PESTL training. Control schools were in the same districts as the treatment schools. The convenience sample was selected based on schools not participating in the professional development with similar characteristics to participating schools and with principals and teachers who were willing to have science experts conduct classroom science teaching observations. The control teachers were observed once during the first year of the study from winter 2008 to spring 2009. To match the treatment group, the control schools were

Table 3.3

Grade Levels and Total Numbers of Participating and Control Teachers

Grade level	Control teachers *N*	PESTL year 1 *n*	PESTL year 2 *n*	PESTL year 3 *n*
Grade 3	1	0	0	0
Grade 4	9	9	9	10
Grade 5	6	10	10	9
Grade 6	4	3	3	3
Total	20	22	22	22

selected based on contextual factors, which included high poverty and diversity characteristics and average school composition.

Control group contamination was addressed by selecting schools that had no teachers who were or had been participants in PESTL. Craven, Marsh, Debus, and Jayasinghe (2001) report that quasi-experimental designs with treatment and control teachers from the same school are subject to “diffusion effects” whereby both experimental and control groups benefit from the intervention, thereby contaminating the control group and biasing evaluations of intervention effects. To limit contamination, diffusion effects were avoided by selecting schools that, as a whole, had not been exposed to the intervention. Control schools were selected based on this premise.

Measures

The data were collected over a 3-year time period (2008-2011). PESTL evaluators made 45-minute classroom observations of science instruction during the regular school day. For triangulation of findings, two observation tools were utilized. Cohen and Manion (2000) defined triangulation as “an attempt to map out, or explain more fully, the richness and complexity of human behavior by studying it from more than one standpoint” (p. 254). Therefore, two tools were used to more fully examine the effect of PESTL professional development on classroom instruction.

The first tool was the Summary Judgment of Science Instruction (see Appendix A for a copy of the tool and the copyright permission to reprint letter), which was adapted from Horizon Research’s Capsule Rating of the Quality of Instruction (Horizon Research, Inc., 2001). The second tool was the PESTL Observation Protocol on Instructional Effectiveness (see Appendix B), which is an original instrument designed for the PESTL program by elementary science experts. Both instruments will be described below, including the establishment of validity and reliability of each.

Summary Judgment of Science Instruction

The Capsule Rating of the Quality of Instruction (see Appendix C) is an observation tool created by Horizon Research, Inc. (HRI) for the *Inside the Classroom* project (HRI, 2001). The *Inside the Classroom* project was coordinated by HRI of Chapel Hill, North Carolina, with support from the National Science Foundation (NSF). Over 30 evaluators in HRI used this tool, which has well-established validity and reliability (Weiss et al., 2003). The reliability of the Capsule Rating of the Quality of Instruction instrument was established by HRI; the Cronbach's alpha reliability for the whole scale was reported at .84 (Weiss et al., 2003). Based on the analyses completed the instrument appeared to be useful for assessing science instruction effectiveness.

The Capsule Rating of Quality of Instruction is a descriptive rubric of science instruction on a five-level rating scale indicating the level of instruction effectiveness. In this final rating of the lesson, the evaluator considers all available information about the lesson, its context and the teacher's purpose, and makes judgment of the relative effectiveness of instruction. The evaluator selects the capsule description that best characterizes the lesson observed (Weiss et al., 2003). The capsule descriptions are as follows.

Level 1: Ineffective Instruction: Two types: (a) passive learning, and (b) activity for activity's sake

Level 2: Elements of Effective Instruction

Level 3: Beginning Stages of Effective Instruction

Level 4: Accomplished, Effective Instruction

Level 5: Exemplary Instruction

This rating is not intended to be an average of all the previous ratings, but should encapsulate the overall assessment of the quality and likely impact of the lesson. The following paragraphs will give a brief description of each level.

Level 1 indicates ineffective instruction. This level would be described as instruction that gave the evaluator little or no evidence of student thinking or engagement with important ideas of science. The evaluator concludes that the instruction is highly unlikely to enhance students' understanding of the discipline or to develop their capacity to successfully "do" science. Ineffective lessons can be characterized as either passive learning or activity for activity's sake. In passive learning the instruction is pedantic and uninspiring (Weiss et al., 2003). Students are passive recipients of information from the teacher or textbook; material is presented in a way that is inaccessible to many of the students. In "activity for activity's sake" students are involved in hands-on activities or other individual or group work, but the lesson lacks a clear sense of purpose and/or a clear link to conceptual development.

Level 2 indicates elements of effective instruction. This level would be described as instruction that contains some elements of effective practice, but there are serious problems in the design, implementation, content, and/or appropriateness for many students in the class. For example, the content may lack importance and/or instruction may not successfully address the difficulties that many students are experiencing. The evaluator concludes that overall the lesson is very limited in its likelihood to enhance students' understanding of the discipline or to develop their capacity to successfully do science.

Level 3 indicates beginning stages of effective instruction. This level would be described as instruction that is purposeful and characterized by quite a few elements of effective practice. Students are, at times, engaged in meaningful work, but there are weaknesses, ranging from substantial to fairly minor, in the design, implementation, or content of instruction. For example, the teacher may short-circuit a planned exploration by telling students what they "should have found," instruction may not adequately address the needs of a number of students, or the classroom culture may limit the accessibility or effectiveness of the lesson. The evaluator concludes that overall the lesson is somewhat limited in its likelihood to enhance students' understanding of the discipline or to develop their capacity to successfully do science.

Level 4 indicates accomplished effective instruction. This level would be described as instruction that is purposeful and engaging for most students. Students actively participate in meaningful work (e.g., investigations, teacher presentations, discussions with each other or the teacher, reading). The lesson is well designed and the teacher implements it well, but adaptation of content or pedagogy in response to student needs and interests is limited. The evaluator concludes that instruction is quite likely to enhance most students' understanding of the discipline and to develop their capacity to successfully do science.

Level 5 indicates exemplary instruction. This level would be described as instruction that is purposeful and all students are highly engaged most or all of the time in meaningful work (e.g., investigation, teacher presentations, discussions with each other or the teacher, reading). The lesson is well designed and artfully implemented, with flexibility and responsiveness to students' needs and interests. The evaluator concludes that instruction is highly likely to enhance most students' understanding of the discipline and to develop their capacity to successfully do science.

The PESTL study slightly adapted the Capsule Rating of the Quality of Instruction tool and renamed it the PESTL Summary Judgment of Effective Instruction. The tool was adapted by including the Utah State Science Core as the guideline for the content mentioned in the original Capsule Rating of the Quality of Instruction tool. The Cronbach's alpha reliability for the adapted PESTL Summary Judgment of Effective Instruction tool used for this study was established at .83. In the adapted tool the description of ineffective instruction in level one remained unchanged; however, in levels two through five, the descriptions include statements about the level at which the instruction aligns with the Utah State Science Core standards and intended learning outcomes (ILO).

In the adapted PESTL Summary Judgment of Effective Instruction, the elements of effective instruction description in level two includes the statement that the lesson content does not align with the Utah State Core. In level three, beginning stages of effective instruction, the description statement includes that the lesson content is aligned with the Core, but the ILO are missing or not featured within the lesson. In Level 4, accomplished effective instruction, the description includes that the lesson is well designed and aligned with the core. In Level 5, exemplary instruction, the description states that the lesson is well aligned with the core and the content and ILO of the core are learned and applied in the lesson. These added statements addressed the need for effective science instruction to align with the science core standards and ILO, which was a major focus of the PESTL professional development program. These adaptions of the Capsule Rating of the Quality of Instruction tool helped make this instrument a more appropriate measure of effective instruction for PESTL participants.

PESTL Observation Protocol on Instructional Effectiveness

The second tool, the PESTL Observation Protocol on Instructional Effectiveness was created specifically to evaluate effective instruction in the areas focused on in the PESTL professional development program. The PESTL Observation Protocol was developed specifically to evaluate effective elementary science instruction as prescribed in *Taking Science to School: Learning and Teaching Science in Grades K-8* (NRC, 2007) and the complementary practitioners book, *Ready, Set, SCIENCE!* (Michaels et al., 2008). A committee of five experts, who were all experienced researchers or teachers in elementary science education, designed the PESTL Observation Protocol on Instructional Effectiveness tool. With this tool, the frequency of effective science teaching practices is measured through a 5-point system, using tally marks and notes. In a pilot study of the first 88 teachers observed in year one of PESTL, the Cronbach's alpha reliability for the

PESTL Observation Protocol on Instructional Effectiveness tool was established at .84.

The PESTL Observation Protocol on Instructional Effectiveness is organized according to the five areas of instructional effectiveness focused on in the PESTL teacher development program. These five instructional areas are: (1) talk and argument, (2) investigation, (3) modeling, (4) science content alignment, and (5) addressing science misconceptions. Within the five areas are thirteen indicators of effective instruction. There are five indicators in the area of talk and argument, three indicators in the area of investigation, three indicators in the area of modeling, one indicator in science content alignment, and one indicator in addressing science misconceptions. Each indicator is scored through tally marks and ratings on a 0-5 point scale.

Talk and argument. Effective teaching indicators within the talk and argument construct include: (a) teacher question to student interaction ratio, (b) number of times teacher actively extends student thinking, (c) number of examples and analogies in presentations, (d) teacher supports relevant inter-student discussion, and (e) teachers' use of accurate science language. Teachers are given tally marks each time an indicator is noted by the evaluator. For example, for indicator (a), the evaluator selects three important science questions the teacher poses and for each question the evaluator tallies the number of student responses up to five. If a new question is posed that appears to be more central to the learning, it is used in lieu of the lower scoring question. The tally mark average of the three questions is calculated for this score. For indicator (b), the evaluator tallies the number of relevant teacher prompts used to extend the discussion specific is to a single question. The evaluator tallies the number of relevant teacher prompts up to five for this score.

For indicator (c), the evaluator observes the number of times the teacher uses examples or analogies to clarify science concepts or principles. The examples and analogies must be connected to the science concept being taught and must help clarify the understanding of science. The total number up to five is the score. For indicator (d), the evaluator looks for the number of times the teacher supports inter-student discussion. For this the evaluator may need to follow the teacher in the classroom or during the discussion to see if the teacher engages the student in giving and receiving of ideas, information and/or discussion from other students. The evaluator totals the number of times the teacher promotes student discussion. The total number up to five is the score.

For indicator (e), the evaluator tallies the number of times the teacher uses accurate science language throughout the science activity. When students make errors in science language, the evaluator looks for the teacher to effectively support the students in using accurate science language. One point is given for each accurate use of a science term up to 5 points. If the teacher presents the misuse of a science term, the score may not exceed one.

Investigation. Effective teaching indicators within the investigation construct include: (f) science investigations are directed by the teacher, (g) science investigations are student centered and in small groups, and (h) science concepts within the investigation are assessed. Teachers are given tally marks each time an indicator is noted by the evaluator. For example, for indicator (f), science investigations are directed by the teacher, the evaluator looks for the teacher to direct students in science investigations.

One point, with a maximum of five, is given for each of the following elements of science investigations as they occur during the learning activity: formulate question, identify observation, formulate a testable question, identify variables as independent, dependent or control, record data, make inferences, discuss limitations of findings, and analyze data. These elements are found in the first ILO in the Utah State Core.

For indicator (g), the evaluator looks for the teacher to direct students to move into small group investigations and to engage in science skills and process. One point, with a maximum of five, is given for each of the following elements of science investigations: formulate question, identify observation, formulate a testable question, identify variable as independent, dependent or control, record data, make inferences, discuss limitations of findings, analyze data. If investigations are conducted as a whole class, the maximum score is three because small group investigations are preferred.

For indicator (h), the evaluator looks for the teacher to assess students' understanding of the investigation using formal or informal assessments. The evaluator also looks for the teacher to pose questions to individuals and the group to clarify and to assess student learning about the science concept within the investigation. Tally marks are given for each formal or informal assessment, including questions, used by the teacher to assess students' understanding. One point, with a maximum of five, is given for each lesson observed.

Modeling and representations. Effective teaching indicators within the modeling and representations construct include: (i) teacher's use of models to demonstrate concepts, (j) teacher's use of models to assess student understanding, and (k) science writing or representations used by students. For indicator (i), the evaluator looks for the teacher to use models to add to students understanding of science concepts or principles in the instruction. Scores are as follows: 1-the models are explained by the teacher only, 3-the models are explained by the student only, and 5-the models are explained by the student and the teacher uses student explanation to extend the learning of others in the classroom. Scores of 2 and 4 are used when the criteria of 3 or 5 are partially met.

For indicator (j), the evaluator looks for the teacher to use students' explanations of the models to assess student knowledge. Scores are as follows: 1-student explanations are inaccurate and the teacher does not use the opportunity for learning, 3-the student explanation is accurate and the teacher accepts it without discussion to extend learning through talk and argument, 4-the teacher requires explanations of the model that are evidenced based, 5-the teacher or students requires evidence and the model is used to extend the science thinking to big ideas and principles. A score of 2 is used when the criteria of 3 is partially met.

For indicator (k), the evaluator looks for the teacher to use students' writing to represent their understanding of science ideas or observations they make. Representations may include diagram, graphs, charts, recording observations in notebook, and so forth. Scores are as follows: 1-no representation used, 2-one representation used, 3-multiple representations are used, 4-the students focus their attention on understanding the phenomena in the representations used, 5-the writing and diagrams provide insight into understanding the science concepts and are used to clarify concepts.

Content. Under content there are two areas of instruction, they are science content alignment and addressing science misconceptions. The indicator of effective teaching within the science content alignment construct is: (l) content is aligned to the Utah state core. The indicator of effective teaching within the addressing science misconceptions construct is: (m) science misconceptions were appropriately addressed. For indicator (l), the evaluator looks for the teacher to have aligned the instruction to the Core. Scores are as follows: 0-instructional objectives are not stated, 1-instructional objectives are stated but not aligned to the Core, 2-instruction targets process skills or ILO only, 3-a clear objective is identify, but the instruction is not well aligned to the objective, 4-a clear objective is well aligned but no ILO is met, 5-a clear objective is identified for instruction and the instruction targets this objective and the ILO supporting the objective.

For indicator (m), the evaluator looks for the teacher to appropriately address student misconceptions during the instruction. Scores are as follows: 0-misconceptions were created or perpetuated by teacher, 1-misconceptions recognized by the teacher but not addressed, 2-student states a misconception, but the teacher does not recognize the misconception, 3-misconceptions were noted and teacher indicated the nature of these misconceptions, 4-teacher corrected misconceptions during the lesson, 5-teacher clarified the misconceptions and provided students with insight through discussion of how to understand the concept.

An overview of the components included in the Summary Judgment of Science Instruction and the PESTL Observation Protocol instruments is given in Table 3.4. These

Table 3.4

PESTL Observation Tools, Level Descriptions, and Sources

Tool #1 Summary Judgment of Science Instruction	Level descriptions & sources

Overall Effectiveness of Science Instruction	• Level 1: Ineffective instruction (2 types)
	  ▪ a) Passive learning
	  ▪ b) Activity for activity sake
	• Level 2: Elements of effective instruction
	• Level 3: Beginning stages of effective instruction
	• Level 4: Accomplished, effective instruction
	• Level 5: Exemplary instruction
	Sources: Horizon Research's Capsule Rating of the Quality of Instruction (HRI, 2001) and Utah State Science Core Curriculum (USOE, 2002).

Tool #2 PESTL Observation Protocol Components	**Indicators & sources of teaching practice descriptions**
Talk and argument	• Nature and frequency of student responses to science questions during classroom discourse • Student use of evidence to support science arguments • Extent to which teacher actively extends student thinking • Use of examples and analogies in presentations • Frequency of relevant inter-student discussion • Precision with which science language is used Designed from descriptions of making thinking visible in *Taking Science to School* (NRC, 2007) and *Ready Set, SCIENCE!* (Michaels et al., 2008).
Investigation & Inquiry	• Science investigations are directed by the teacher • Science investigations are student centered and in small groups • Science investigations are aligned to science concepts • Science concepts within the investigation are assessed Designed from descriptions of science investigations in *Taking Science to School* (NRC, 2007) and *Ready Set, SCIENCE!* (Michaels et al., 2008).
Modeling & Representations	• Frequency of the use of models to teach concepts • Use of models to assess student understanding • Use of science writing or representations by students Designed from descriptions of modeling in *Taking Science to School* (NRC, 2007) and *Ready Set, SCIENCE!* (Michaels et al., 2008).
Science Content Alignment	• Instructional objectives are aligned to the state core curriculum content • Instruction targets science practices described in the State Core Curriculum "Intended Learning Outcomes" Source: Utah State Science Core Curriculum (USOE, 2002).
Addressing Science Misconceptions	• Science misconceptions were appropriately addressed Designed from descriptions of conceptual change in *Taking Science to School* (NRC, 2007) and *Ready Set, SCIENCE!* (Michaels et al., 2008).

two tools were used to gather data for this study. The procedures of this study will be described in the following section.

Procedures

Data used for this study was pre-existing. Data were collected during classroom observations of 22 cluster sampled treatment teachers who were evaluated each year over the 3 years of professional development. To establish a baseline for nontreatment, 20 convenience-sampled control teachers were observed once during the first year of the study. PESTL evaluators made 45-minute classroom observations of science instruction during the regular school day using two instruments: the PESTL Observation Protocol on Instructional Effectiveness and the Summary Judgment of Science Instruction. Data were collected over a 3-year time period from Winter 2008 to Spring 2011. This researcher began to study the PESTL program in 2010 and was involved as an evaluator through the last full year of the professional development program.

This researcher examined the literature, specifically standards documents and research related to elementary science education reform. The specific characteristics sought after for this study included a long-term professional development program in elementary science education that developed partnerships between science experts and elementary teachers and that promoted scientific discourse and inquiry in the classroom. The criteria sought after for analysis in this study included a program that could be examined through quantitative data that measured changes in elementary science instruction over time. The PESTL program met the characteristics and criteria this researcher desired to study. PESTL aligns with the current literature on science education reform; it is a long-term professional development program designed to improve elementary science education through partnerships and teacher science literacy. PESTL not only promotes scientific discourse and inquiry in the classroom, it also provides professional development in many additional components of science education. Therefore, PESTL was determined to be the ideal program to examine for this study.

Science education experts designed the PESTL professional development program to address the recommendations for elementary science education reform. Based on these recommendations, the PESTL program was designed to be a sustained professional development program that lasted for 3 year. From across the state of Utah, 148 third- thru sixth-grade teachers were selected to receive the sustained professional development. Teacher selection occurred in the spring of 2008. PESTL intervention commenced in the summer of 2008.

Six rater/evaluators, who were all experienced researchers and experts in elementary science education, gathered data for this study. To establish reliability and inter-rater agreement measures, program coordinators provided 2-day trainings for all evaluators before they were able to make classroom observations. Evaluators were retrained each year to ensure fidelity of rater agreement over time. Training sessions included education on the tools and the constructs being measured. Following the training, evaluators made three parallel observations with two evaluators observing the same lesson. The program coordinators checked inter-rater reliability following each observation. Debrief sessions followed the training sessions, where the trainee evaluator could provide defense for his or her scores. Evaluators were also required to attend portions of the PESTL professional development summer sessions.

This researcher was trained as an evaluator for the third year of the program and conducted treatment and control observations during year three of PESTL. The researcher attended the full summer session in 2010 and received the PESTL evaluator training in December of 2010. The researcher also observed all other components of the PESTL program through the 2010-11 academic year. The following procedures were carried out in order to gather the data for this study.

Evaluating Classroom Instruction of Treatment Teachers

PESTL evaluators conducted classroom observations of the 22 treatment teachers. Each teacher was observed three times, once each year of treatment. Observations were made in the spring, near the end of each year of professional development. Other PESTL participants were observed for continued program funding and evaluation purposes: however, only the 22 participants for this study were cluster sampled and followed throughout the 3 years of the PESTL program. Observation appointments for treatment teachers were arranged through the school PESTL PLC facilitator.

Evaluating Classroom Instruction of Nonparticipants

To establish a baseline for current non-participant practice, the PESTL evaluators observed and rated 20 control teachers using the same two instruments utilized for the treatment group observations. Control teachers were observed once. The 20 control observations were made during the first year of the program Control observations were made during the winter of 2008 and spring of 2009. Times of visitation were arranged through the school principal. Control observation scores on both instruments were analyzed using independent *t* tests for equality of means.

Sustained Intervention

Participants received sustained intervention for all 3 years of PESTL professional development. Unlike traditional “one-time” workshops, sustained professional development provides ongoing information and support to teachers over an extended period of time. Sustained professional development can extend over a few months or even over multiple years (Darling-Hammond & McLaughlin, 1995). Over the course of the program, very few teachers opted to no longer participate while other teachers chose to enroll after year one. The 22 participants selected for analysis in this data set completed the entire 3 years of the program.

Designing this Study

The researcher designed this study to examine the effects of sustained professional development on classroom instruction of elementary school teachers over time. PESTL data collection was not originally designed to employ repeated measures analysis, so this researcher recommended the selection of a random sample of treatment teachers observed during year one and two to be observed again in year three. This formed the fundamental design of this study.

Evaluating Classroom Instruction of Treatment Group in Third Year of Intervention

The researcher identified a sample of 22 treatment teachers for this study. These teachers had been random cluster sampled (Lohr, 1999) for observations in years one and two. Trained evaluators including this researcher conducted the third and final classroom observations of these 22 teachers near the end of year three of intervention. An overview of the PESTL treatment and observations for each year is illustrated in Table 3.5.

Analyzing the Data

Two data sets were analyzed. One set was data gathered from the Summary of Effective Science Instruction tool. The other set was data gathered from the PESTL Observation Protocol tool. Data analyses were independent *t* test and one-way repeated

Table 3.5

Overview of PESTL Sustained Treatment and Participant Observations

Year	Sustained treatment	Observations
1	Summer seminar - 5 full days School PLCs - 8 times Grade-level alignments - 2 times (Fall and Spring) Fall content course Mid-winter institute	Ten schools were randomly selected for cluster sample. 22 sample teachers were observed. 20 control teachers were observed.
2	Summer seminar - 5 full days School PLCs - 6 times Grade-level alignments - 2 times (Fall and Spring) Fall content course Mid-winter institute	22 sample teachers were observed.
3	Summer seminar - 5 full days School PLCs - 4 times Grade-level alignments - 3 times (Fall, Winter, and Spring) Fall content course Mid-winter institute	22 sample teachers were observed.

measures ANOVAs. All data from the classroom observations were analyzed using Statistical Package for Social Sciences (SPSS) for Apple (OSX). The research questions and corresponding hypotheses are listed below.

Research Question One

Research question one asked, “Does sustained professional development of 3 years, which was designed to meet recommendations for reform in science education, transform overall classroom science instruction of elementary school teachers?” The hypothesis is that third- through sixth-grade teachers who participated in sustained professional development in elementary science education would display improvements in overall science instruction over a 3-year time period.

Research Question Two

Research question two asked, “Does sustained professional development of 3 years transform the instructional practice of elementary school teachers in *specific components* of reformed science education?” The hypothesis is that third through sixth grade teachers who participated in sustained professional development in elementary science education will display improvements in specific components of reformed science education over a 3-year time period. To determine the effects of sustained professional development on the reformed instructional models provided in the professional development the five constructs of the PESTL Observation Protocol were compared separately. Five subquestions for question two were identified as follows.

2a. Does sustained professional development affect the practice of classroom discussion and argumentation in the science instruction of elementary school teachers?

2b. Does sustained professional development affect the practice of scientific investigation in the classroom science instruction of elementary school teachers?

2c. Does sustained professional development affect the practice of teacher and student modeling in the classroom science instruction of elementary school teachers?

2d. Does sustained professional development affect the practice of aligning science instruction to core standards in the classroom instruction of elementary school teachers?

2e. Does sustained professional development affect the practice of appropriately addressing student science misconceptions in the classroom instruction of elementary school teachers?

Research Question Three

Research question three asked, "How does sustained professional development of 3 years transform overall classroom science instruction and the specific components of reformed science instruction of elementary school teachers over time?" The hypothesis is that third through sixth-grade teachers who participate in sustained professional development in elementary science education will demonstrate patterns of change in instructional practices over time. It is expected that unfamiliar or more complicated components of instruction will take a greater amount of time to improve. It is also expected that overall instruction scores may improve before specific components of instruction do. Examining patterns of change within facets of instruction can provide important information about sustained professional development as well as areas of focus for future study. Four subquestions for question three were identified as follows.

3a. What components of reformed science instruction are most influenced by sustained professional development over time?

3b. What components of reformed science instruction are least influenced by sustained professional development over time?

3c. Are there patterns of change among the components of reformed science instruction of elementary school teachers over time?

3d. Are there interactions between performance in overall classroom science instruction and the practice of specific components of reformed science instruction of elementary school teachers over time?

The researcher prepared a full study and description of the research questions by reporting on how the sustained teacher development of PESTL in elementary science education influenced classroom instruction over a 3-year time period. The researcher also prepared a full study and description of the findings relating to the hypotheses that third through sixth grade teachers who participate in sustained teacher development in elementary science education will display improvements in overall science instruction, will display improvements in specific components of reformed science education, and will demonstrate patterns of change in instructional practices over a 3-year time period. These results are found in the next chapter.

CHAPTER 4

FINDINGS

The science expert completed the evaluation of Ms. King's classroom instruction and sent the observation scores to the PESTL program data record keeper. The record keeper then entered Ms. King's Summary of Effective Science Instruction rating and PESTL Observation Protocol scores into an Excel file, along with the other observation scores of participating and control teachers from across the state. After 3 years of sustained professional development and observing the participants, the data were ready to be analyzed. The researcher determined the proper data sets and analyses to use in order to examine the effects of the sustained professional development program on classroom instruction. The results of these analyses would be examined for significant difference between groups and changes over time. The findings would reveal the effects of sustained professional development on the classroom science instruction of elementary school teachers.

Introduction

The purpose of this study was to determine the extent to which sustained professional development in elementary science education affects classroom instruction over a 3-year period. Study participants participated in 3 years of PESTL professional development, which was designed to meet reforms in elementary science education. In order to determine the changes in classroom instruction two data sets were analyzed. The first data set were ratings from the Summary of Effective Science Instruction tool. The second data set were ratings from the PESTL Observation Protocol tool. With the PESTL Observation Protocol data, five different areas of effective science instruction were analyzed. The five areas were: (a) talk and argument, (b) investigation, (c) modeling, (d) content alignment, and (e) addressing misconceptions.

Data Analysis

Data analyses were independent *t* tests and one-way ANOVAs. Independent *t* tests were used for examining data between the control group and treatment group. The researcher chose to use *t* tests rather than ANOVAs because the control teachers were not the same as the treatment teachers; therefore, the control scores did not qualify for repeated measures analyses. A Bonferroni correction was made for the *p*-value significance level in order to avoid a spurious positive and to account for the number of comparisons being performed. For the Bonferroni correction, the alpha values were lowered to $\alpha = .0125$. One-way repeated measures ANOVAs with Bonferroni post-hoc were used for examining change over time within the treatment group. All data from the classroom observations were analyzed using Statistical Package for Social Sciences (SPSS) for Apple (OSX).

The reliability of the scales for the Summary Judgment of Effective Science Instruction and the PESTL Observation Protocol was established. Cronbach's alpha on the two tools and five PESTL subscales reflects high levels of internal consistency: Summary Judgment of Effective Science Instruction $\alpha = .83$, PESTL Observation Protocol total $\alpha = .84$, talk and argument $\alpha = .87$, investigation $\alpha = .87$, and modeling $\alpha = .84$. The Cronbach's alpha reliability for the entire observation scale was reported at .84, which is high and indicates strong internal consistency among the measures.

Normality of the two observation tools was checked by looking at the skewness and kurtosis values of both tools overall and each PESTL subscale. Skewness is a measure of symmetry. It checks that a data set is normally distributed to the left and right of the mean. Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution. The acceptable range for both skewness and kurtosis is within 2.0 and -2.0. The skewness for the Summary Judgment of Effective Science Instruction was -.77 with kurtosis at .02. Skewness for PESTL Observation Protocol total was -.08 with kurtosis at -1.28. The PESTL subscale normality included talk and argument -.74 skewness with -.56 kurtosis, investigation -1.04 skewness with .53 kurtosis, modeling -.37 skewness with -.04 kurtosis, content alignment -1.75 skewness with 3.02 kurtosis, and addressing misconception -.29 skewness with -1.26 kurtosis.

All but one of subscales fell within the acceptable range and reflected a relatively normal distribution. The component of content alignment with 3.02 kurtosis fell outside the acceptable range of normality. The content alignment scores have a distinct peak near the mean because the entire set of scores for both treatment and control teachers were high. There was less spread in these scores, because, as this study found, most elementary teachers will align their instruction to the core objectives, especially when they are being observed. This caused the content alignment subscale to have a higher kurtosis measure. The analysis of this particular subscale provided findings that will help improve weaknesses in the PESTL Observation Protocol tool and professional development program. Because of the valuable information this data set provided, it was retained and analyzed with the same parametric statistics as the other subscales.

Research Question One

Research question one asked, “Does sustained professional development of 3 year, which was designed to meet recommendations for reform in science education, transform overall classroom science instruction of elementary school teachers?" The overall science instruction of participating teachers was measured each year through classroom observations using the Summary of Effective Science Instruction tool and the PESTL Observation Protocol total score.

Overall Science Instruction

The overall science instruction rating for each lesson was evaluated using the Summary Judgment of Science Instruction (out of 5), and the PESTL Observation Protocol total score (out of 65). The Summary Judgment of Science Instruction rating ranged from level 1 = *Ineffective Instruction* to level 5 = *Exemplary Instruction*, based on the observer's judgment of how likely the lesson was to enhance most students' understanding of the discipline and to develop their capacity to successfully do science. The following is a brief description of each level:

Level 1: Ineffective Instruction—two types: (a) passive learning, and (b) activity for activity's sake

Level 2: Elements of Effective Instruction

Level 3: Beginning Stages of Effective Instruction

Level 4: Accomplished, Effective Instruction

Level 5: Exemplary Instruction

Level 1 indicates ineffective instruction. This level would be described as instruction that gave the evaluator little or no evidence of student thinking or engagement with important ideas of science. Level 2 indicates elements of effective instruction. This level would be described as instruction that contains some elements of effective practice, but there are serious problems in the design, implementation, content, and/or appropriateness for many students in the class. Level 3 indicates beginning stages of effective instruction. This level would be described as instruction that is purposeful and characterized by quite a few elements of effective practice. Level 4 indicates accomplished effective instruction. This level would be described as instruction that is purposeful and engaging for most students. Level 5 indicates exemplary instruction. This level would be described as instruction that is purposeful and all students are highly engaged most or all of the time in meaningful work.

The PESTL Observation Protocol total was the sum of thirteen indicator scores from the five different components of reformed science instruction included in the instrument. The five components were:

Component 1: Talk and argument

Component 2: Investigation

Component 3: Modeling

Component 4: Science Content Alignment

Component 5: Addressing Science Misconceptions

The five effective teaching indicators within the talk and argument component include: (a) teacher question to student interaction ratio, (b) number of times teacher actively extends student thinking, (c) number of examples and analogies in presentations, (d) teacher supports relevant interstudent discussion, and (e) teachers' use of accurate science language.

The three effective teaching indicators within the investigation component include: (f) science investigations are directed by the teacher, (g) science investigations are student centered and in small groups, and (h) science concepts within the investigation are assessed. The three effective teaching indicators within the Modeling component include: (i) teacher's use of models to demonstrate concepts, (j) teacher's use of models to assess student understanding, and (k) science writing or representations used by students.

There are two components that only had one indicator each, content alignment and addressing misconceptions. The indicator of effective teaching within the science content alignment component is: (l) lesson content is aligned to the Utah state core, and the indicator of effective teaching within the addressing science misconceptions component is: (m) science misconceptions were appropriately addressed by the teacher. These indicators (a-m) comprise the PESTL Total score.

The descriptive statistics for the Summary Judgment and PESTL Observation Protocol are provided in Table 4.1. The average mean scores and standard deviations of each group are represented.

To determine the effects of sustained professional development on overall classroom instruction ratings, independent *t* tests were used to compare the control group scores to year one of treatment scores. The researcher chose to use *t* tests rather than ANOVAs because the control teachers were not the same as the treatment teachers; therefore, the control scores did not qualify for repeated measures analyses. A Bonferroni correction was made for the *p*-value significance level in order to avoid a spurious

Table 4.1

Means and Standard Deviations for Summary Judgment and PESTL Observation Protocol

Variables	*n*	*M*	*SD*
Summary judgment			
Control	20	2.20	1.11
Treatment Year 1	22	2.73	1.12
Treatment Year 2	22	3.41	1.05
Treatment Year 3	22	4.14	0.64
PESTL protocol			
Control	20	25.45	14.43
Treatment Year 1	22	32.64	11.52
Treatment Year 2	22	47.09	12.11
Treatment Year 3	22	57.41	4.76

positive and to account for the number of comparisons being performed. For the Bonferroni correction the alpha values were lowered to $\alpha = .0125$.

The independent categorical variable was group and the dependent continuous variable was the Summary Judgment of Science Instruction and PESTL Observation Protocol Total score. Separate analyses were conducted for each variable. Table 4.2 reports the results of the independent *t* test. The average mean scores, standard deviations, mean differences, and significance *p*-values of each group are represented. A summary statement follows.

While mean scores were higher in both measures for the treatment group year one of the professional development program, there were no significant differences between

Table 4.2

Mean Difference in Overall Ratings by Experimental Condition for Year One

	Control group (n = 20)		Treatment year one (n = 22)			
Variable	*M*	*SD*	*M*	*SD*	Mean difference	*p*
Summary judgment	2.20	1.11	2.73	1.12	.53	.133**
PESTL total	25.45	14.43	32.64	11.52	7.19	.081**

** Mean difference is not significant at the Bonferroni corrected .0125 level.

the control group and treatment group year one. The treatment group had higher Summary Judgment of Science Instruction scores (M = 2.73) after 1 year when compared to the control teachers (M = 2.20); however, these differences were not statistically significant (p = 0.13). Likewise, the treatment group had higher PESTL Observation Protocol Total scores (M = 32.64) after 1 year when compared to the control teachers (M = 25.45), but again these differences were not statistically significant (p = .085). The results suggest that teachers who participated in PESTL professional development for only 1 year did not have significantly higher overall scores in science instruction when compared to teachers who did not participate in the PESTL professional development. This is likely due to the need for sustained professional development in order to significantly improve elementary science instruction.

One-way repeated measures ANOVAs were conducted to examine differences in the treatment group over time. The independent categorical variable was year of treatment and the dependent continuous variable was the Summary Judgment of Science Instruction and PESTL Observation Protocol Total. Separate analyses were conducted for each variable. The mean difference scores were computed to analyze significant differences in ratings over time. Table 4.3 reports the finding of the repeated measures ANOVA with the Summary Judgment of Effective Science Instruction and PESTL Observation Protocol Total data sets.

For the Summary Judgment of Effective Science Instruction the repeated measures ANOVA revealed significant differences between the groups, F (2, 63) = 11.81, p = .000. A Bonferroni post hoc analysis was conducted to examine differences between the groups. The post hoc showed that the treatment group year three had significantly higher mean scores than year one (p = .000) and year two (p = .044). This suggests that participating in sustained professional development over 3 years is more likely to

Table 4.3

Year-by-Year Means and Significances of Overall Measures for the Treatment Group All 3 Years

Variable	Treatment group *n = 22*	Comparison year	Mean difference	*p*
Summary Judgment	Year 1	Year 2	-0.68	.066**
		Year 3	-1.41	.000*
	Year 2	Year 1	0.68	.066**
		Year 3	-0.73	.044*
	Year 3	Year 1	1.41	.000*
		Year 2	0.73	.044*
PESTL Total Score	Year 1	Year 2	-14.46	.000*
		Year 3	-24.77	.000*
	Year 2	Year 1	14.46	.000*
		Year 3	-10.32	.003*
	Year 3	Year 1	24.77	.000*
		Year 2	10.32	.003*

* Mean difference is significant at the 0.05 level.
** Mean difference is not significant at the 0.05 level.

increase effectiveness of the classroom science instruction of elementary school teachers.

The post hoc analysis revealed no significant difference between those who participated in 1 year of professional development training and those who participated in 2 years of professional development training (p = .066). The mean score was higher after 2 years (M = 3.41) when compared to 1 year (M = 2.73); however, these differences were not statistically significant. This may be due to the small sample size of treatment teachers. This result does indicate that teachers who continue into 3 years of professional development are more likely to increase their effective science instructional practices over and above those who participated in just 1 or 2 years of professional development.

For the PESTL Observation Protocol Total score the repeated measures ANOVA revealed significant differences between the groups, F (2, 63) = 33.82, p = .000. Post hoc analysis was conducted to examine differences between the groups. The post hoc analysis revealed that the treatment group year three had significantly higher mean scores than year one (p = .000) and year two (p = .000). The analysis further revealed significant differences between teachers who participated in 2 years of professional development training and those who participated in 3 years of professional development training (p = .003). Unlike the Summary scores, which did not reveal significant improvement from year one to two, the PESTL Total scores revealed significant improvements each year. The PESTL Observation Protocol was designed specifically to measure the practices of instruction focused on in the PESTL program. This may explain the greater increase in this measure. This result does indicate that participating in sustained professional development is likely to increase the effective science instructional practices of elementary school teachers over time.

Indeed, the treatment teacher scores after 3 years of professional development were significantly higher than treatment teacher scores after 1 year and 2 years of professional development. These findings indicate that teachers who continue into 3 years of development are more likely to increase their practice of effective science instruction over and above those who participated in just 1 or 2 years of professional development.

The findings suggest that participating in sustained professional development does increase the overall classroom science instruction of elementary school teachers. These finding support the hypothesis that teachers who receive 3 years of professional development are more likely to have enhanced overall science instructional practices. It appears there are significant differences between groups. The treatment teacher scores after 3 years of professional development were significantly higher than treatment teacher scores after 1 year and after 2 years of professional development.

Research Question Two

Research question two asked, "Does sustained professional development of 3 years transform the instructional practice of elementary school teachers in specific components of reformed science education?" The researcher analyzed the data gathered from the PESTL Observation Protocol tool to answer this question.

To determine the effects of sustained professional development on the reformed instructional models provided in the professional development the five components of the PESTL Observation Protocol were compared separately. Five subquestions for question two were identified as follows.

2a. Does sustained professional development affect the practice of classroom discussion and argumentation in the science instruction of elementary school teachers?

2b. Does sustained professional development affect the practice of scientific investigation in the classroom science instruction of elementary school teachers?

2c. Does sustained professional development affect the practice of teacher and student modeling in the classroom science instruction of elementary school teachers?

2d. Does sustained professional development affect the practice of aligning science instruction to core standards in the classroom instruction of elementary school teachers?

2e. Does sustained professional development affect the practice of appropriately addressing student science misconceptions in the classroom instruction of elementary school teachers?

Independent *t* tests were used to compare the control group scores to year one of treatment scores. The independent categorical variable was group and the dependent continuous variables were the PESTL components as related to each subquestion. PESTL component scores from observations from year one of treatment were compared to the PESTL component scores from the control group.

In the one-way repeated measures ANOVAs the independent categorical variable was year of treatment and the dependent continuous variables were the PESTL components as related to each subquestion. PESTL component scores from observations of the treatment group from years one, two, and three were compared over time.

The descriptive statistics for the five components of reformed science instruction included in the PESTL Observation Protocol are provided in Table 4.4. The average mean scores and standard deviations of each group are represented.

Table 4.4

Means and Standard Deviations for Components of Reformed Science Education

Group	*n*	*M*	*SD*
Talk and argument (out of 25)			
Control	20	11.40	5.90
Treatment year 1	22	12.41	4.42
Treatment year 2	22	18.91	4.59
Treatment year 3	22	23.05	1.76
Investigation (out of 15)			
Control	20	4.65	3.53
Treatment year 1	22	6.64	4.18
Treatment year 2	22	9.68	3.14
Treatment year 3	22	11.95	1.65
Modeling (out of 15)			
Control	20	4.00	3.93
Treatment year 1	22	6.68	3.77
Treatment year 2	22	8.77	3.52
Treatment year 3	22	10.36	3.00
Content alignment (out of 5)			
Control	20	3.15	1.79
Treatment year 1	22	3.55	0.74
Treatment year 2	22	4.45	1.22
Treatment year 3	22	4.77	1.07
Misconceptions (out of 5)			
Control	20	2.25	1.65
Treatment year 1	22	2.36	1.18
Treatment year 2	22	3.27	1.52
Treatment year 3	22	4.27	1.24

Talk and Argument

The talk and argument component consists of five indicators. Each is rated on a 5-point scale. They are:

- Student/teacher questioning and interaction ratio
- Number of times teacher actively extends student thinking
- Number of examples and analogies in presentations
- Teacher supports relevant inter-student discussion
- Teacher uses accurate science language.

Teachers were given tally marks each time an indicator was noted by the evaluator. A full description of the PESTL Observation Protocol rating scale is included in Appendix B.

To address question two and specifically subquestion 2a, an independent *t* test was used to determine differences between the control and treatment year one. The researcher used *t* tests rather than ANOVAs because the control teachers were not the same as the treatment teachers, therefore the control scores did not qualify for repeated measures analyses. A Bonferroni correction was made for the *p*-value significance level in order to avoid a spurious positive and to account for the number of comparisons being performed. For the Bonferroni correction the alpha values were lowered to $\alpha = .0125$.

In the independent *t* test the independent categorical variable was group and the dependent continuous variable was the talk and argument score. Talk and argument scores from observations from treatment year one were compared to the Talk and argument scores from the control group.

There was no significant difference between the control group and treatment year one $t(40) = -.631$, $p = .531$. The treatment group had higher talk and argument mean scores ($M = 12.41$) after 1 year compared to the control teachers ($M = 11.40$); however, these differences were not statistically significant ($p = .531$). This may be due to the small sample size of treatment teachers or may indicate the professional development year one was not effective in changing practice. This finding does indicate that teachers who participate in PESTL need two or more years of professional development to significantly increase their effective science instructional practices of Talk and argument. The results are that teachers who participate in PESTL professional development do not have significantly higher scores in talk and argument after 1 year when compared to a control group of nonparticipating teachers.

In the one-way repeated measures ANOVAs the independent categorical variable was year of treatment and the dependent continuous variable was the Talk and argument score. Talk and argument scores from observations of the treatment group from year one, year two, and year three were compared over time. Table 4.5 reports the findings of the repeated measures ANOVA.

Table 4.5

Repeated Measures ANOVA for Talk and argument of Treatment Group Over 3 Years

Group	Comparison group	Mean difference	*p*
Treatment year 1	Treatment year 2	-6.500	.000*
	Treatment year 3	-10.636	.000*
Treatment year 2	Treatment year 1	6.500	.000*
	Treatment year 3	-4.136	.002*
Treatment year 3	Treatment year 1	10.636	.000*
	Treatment year 2	4.136	.002*

* Mean difference is significant at the 0.05 level.

The ANOVA revealed significant differences between the groups, $F(2, 63) = 43.495$, $p = .000$. A post hoc analysis, using Bonferroni correction, was used to examine differences between the groups. The post hoc showed that the treatment group year three had significantly higher mean scores than at year one ($p = .000$) and at year two ($p = .002$). This suggests that participating in sustained professional development may increase the use of Talk and argument in the classroom science instruction of elementary school teachers. The post hoc analysis also revealed that the treatment group year two had significantly higher mean scores than at year one ($p = .000$). This finding supports the hypothesis that teachers who receive two or more years of professional development are more likely to have enhanced science instructional practices in Talk and argument.

Investigation

There are three observational items included in the investigation construct, each rated on a 5-point scale. They are:

- Science investigations are directed by the teacher
- Science investigations are student centered and in small groups
- Science concepts within the investigation are assessed

Teachers were given tally marks each time an indicator was noted by the evaluator. A full description of the PESTL Observation Protocol rating scale is included in Appendix B.

To address question two and specifically subquestion 2b, an independent *t* test was used to determine differences between the control and treatment year one. The researcher chose to use *t* tests rather than ANOVAs because the control teachers were not the same as the treatment teachers; therefore, the control scores did not qualify for repeated measures analyses. A Bonferroni correction was made for the *p*-value significance level in order to avoid a spurious positive and to account for the number of comparisons being performed. For the Bonferroni correction the alpha values were lowered to $\alpha = .0125$.

In the independent *t* tests the independent categorical variable was group and the dependent continuous variable was the Investigation score. Investigation scores from observations from year one of treatment were compared to the investigation scores from the control group.

There was no significant difference between the control group and treatment year one $t(40) = -1.66$, $p = .106$. The treatment group had higher Investigation mean scores ($M = 6.64$) after 1 year compared to the control teachers ($M = 4.65$); however, these differences were not statistically significant ($p = .106$). This may be due to the small sample size of treatment teachers or may indicate the professional development year one was not effective in changing practice. This finding suggests that teachers need to continue with two or more years of professional development in order to significantly improve their practice of Investigation. Results suggest that teachers who participate in PESTL professional development do not have significantly higher scores in Investigation after 1 year when compared to a control group of nonparticipating teachers.

In the one-way repeated measures ANOVAs, the independent categorical variable was year of treatment and the dependent continuous variable was the investigation score. Investigation scores from observations of the treatment group from years one, two, three were compared over time. Table 4.6 reports finding of the repeated measures ANOVA.

Table 4.6

Repeated Measures ANOVA for Investigation of Treatment Group Over 3 Years

Group	Comparison year	Mean difference	*p*
Treatment year 1	Treatment year 2	-3.045*	.007*
	Treatment year 3	-5.318*	.000*
Treatment year 2	Treatment year 1	3.045*	.007*
	Treatment year 3	-2.273	.061**
Treatment year 3	Treatment year 1	5.318*	.000*
	Treatment year 2	2.273	.061**

* Mean difference is significant at the 0.05 level.
** Mean difference is not significant at the 0.05 level.

The ANOVA revealed significant differences between the groups, F (2, 63) = 15.65, p = .000. A Bonferroni post hoc analysis was used to examine differences between the groups. The post hoc showed that the treatment group year three had significantly higher mean scores than year one (p = .000). The post hoc analysis also revealed a significant difference between those who participated in 1 year of training and 2 years of training (p = .007). This suggests that participating in sustained professional development does increase the use of Investigation in the classroom science instruction of elementary school teachers.

The post hoc analysis revealed no significant difference between those who participated in 3 years of professional development training and those who participated in 2 years of professional development training (p = .061). The mean score was higher after 3 years (M = 11.95) when compared to 2 years (M = 9.68); however, these differences were not statistically significant. The participant scores improved over time but not significantly from year two to year three. This may be due to the small sample size of treatment teachers. These results support the hypothesis that teachers who receive additional professional development are more likely to have enhanced science instructional practices in Investigation.

Modeling

The modeling construct consists of three observational items, each rated on a 5-point scale. They are:

- Teacher uses models to demonstrate concepts
- Teacher uses models to assess student understanding
- Science writing or representations are used by students

Teachers were given scores for each indicator as outlined on the observation tool and noted by the evaluator. A full description of the PESTL Observation Protocol rating scale is included in Appendix B.

To address question two and specifically subquestion 2c, an independent *t* test was used to determine differences between the control and treatment year one. The researcher chose to use *t* tests rather than ANOVAs because the control teachers were not the same as the treatment teachers, therefore the control scores did not qualify for repeated measures analyses. A Bonferroni correction was made for the *p*-value significance level in order to avoid a spurious positive and to account for the number of comparisons being performed. For the Bonferroni correction the alpha values were lowered to $\alpha = .0125$. In the independent *t* tests, the independent categorical variable was group and the dependent continuous variable was the modeling score. Modeling scores from observations from year one of treatment were compared to modeling scores from the control group.

There were differences between the control group and treatment year one t(40) = -2.26, p = .030; however, with the Bonferroni correction this significance was considered possibly spurious and therefore is disregarded. The treatment group had higher Modeling scores (M = 6.68) after 1 year when compared to the control teachers (M = 4.00). The results suggest that teachers who participate in PESTL professional development have higher scores in Modeling after 1 year of professional development but these differences may not be enough to suggest significant improvement.

In contrast to the other components, which showed no significant difference between the control and year one of treatment at the 0.05 level, Modeling did demonstrate significant difference before the Bonferroni correction. This difference may be due to the lack of background elementary teachers have with the concept of modeling in science instruction. The use of Modeling rating of the control teachers was very low. The professional development in year one may provide enough exposure to the concept and practice of Modeling to explain the differences at the 0.05 level between groups but does not meet the α = .0125 level.

A one-way repeated measures ANOVA was used to examine changes over time in the practice of Modeling. The independent categorical variable was year of treatment and the dependent continuous variable was the Modeling score. Modeling scores from observations of the treatment group from year one, year two, and year three. Table 4.7 reports the finding of the repeated measures ANOVA.

Table 4.7

Repeated Measures ANOVA for Modeling of Treatment Group Over 3 Years

Group	Comparison group	Mean difference	*p*
Treatment year 1	Treatment year 2	-2.091	.145**
	Treatment year 3	-3.682*	.002*
Treatment year 2	Treatment year 1	2.091	.145**
	Treatment year 3	-1.591	.392**
Treatment year 3	Treatment year 1	3.682*	.002*
	Treatment year 2	1.591	.392**

*Mean difference is significant at the 0.05 level.
**Mean difference is not significant at the 0.05 level.

The repeated measures ANOVA revealed significant differences between the groups, F (2, 63) = 6.320, p = .003. A post hoc analysis was used to examine differences between the groups. The post hoc showed that the treatment group year three had significantly higher mean scores than at year one (p = .002). The post hoc analysis revealed no significant difference between those who participated in 3 years of professional development training and those who participated in 2 years of professional development training. The mean score was higher at 3 years (M = 10.36) when compared at 2 years (M = 8.77); however, these differences were not statistically significant (p = .392).

Furthermore, the post hoc revealed no significant difference between those who participated in 2 years of professional development training and those who participated in 1 year of professional development training. The mean score was higher at 2 years (M = 8.77) when compared at 1 year (M = 6.68); however, these differences were not statistically significant (*p* = .145). These findings suggest that participant scores improved over time but not significantly from year one to year two, nor from year two to year three. This may indicate the professional development for the first 2 years was not effective in changing the practice of Modeling; however it may also suggest that teachers need sustained professional development of 3 years in order to build understanding and practices of Modeling at the level needed to successfully implement it into classroom instruction.

Science Content Alignment

The science content alignment construct consists of one observational item rated on a 5-point scale, which is:

- Content is aligned to the core

Teachers were given a score for this indicator as outlined on the observation tool and noted by the evaluator. A full description of the PESTL Observation Protocol rating scale is included in Appendix B.

To address question two and specifically subquestion 2d, an independent *t* test was used to determine differences between the control group and treatment year one. The researcher chose to use *t* tests rather than ANOVAs because the control teachers were not the same as the treatment teachers, therefore the control scores did not qualify for repeated measures analyses. A Bonferroni correction was made for the *p*-value significance level in order to avoid a spurious positive and to account for the number of comparisons being performed. For the Bonferroni correction the alpha values were lowered to $\alpha = .0125$.

In the independent *t* tests, the independent variable was group and the dependent continuous variable was the content alignment score. Content alignment scores from observations from year one of treatment were compared to the content alignment scores from the control group.

There was no significant difference between the control group and treatment year one $t(40) = -0.95$, $p = .346$. The treatment group had higher content alignment mean scores ($M = 3.55$) after 1 year compared to the control teachers ($M = 3.15$); however, these differences were not statistically significant ($p = .346$). This may be due to the small sample sizes or may indicate the professional development year one was not effective in changing practice. This finding does indicate that teachers who participate in PESTL need two or more years of professional development to significantly increase their effective science instructional practices of content alignment. The results are that teachers who participate in PESTL professional development do not have significantly higher scores in content alignment after 1 year when compared to a control group of nonparticipating teachers.

In the one-way repeated measures ANOVAs the independent categorical variable was year of treatment and the dependent continuous variable was the Content Alignment score. Content Alignment scores from observations of the treatment group from year one, year two, and year three were compared. Table 4.8 reports the finding of the repeated measures ANOVA.

The ANOVA revealed significant differences between the groups, $F(2, 63) = 8.420$, $p = .001$. A post hoc analysis with Bonferroni was used to examine differences

Table 4.8

Repeated Measures ANOVA for Content Alignment of Treatment Group Over 3 Years

Group	Comparison year	Mean difference	*p*
Treatment year 1	Treatment year 2	-.909	.014*
	Treatment year 3	-1.227	.001*
Treatment year 2	Treatment year 1	.909	.014*
	Treatment year 3	-.318	.928**
Treatment year 3	Treatment year 1	1.227	.001*
	Treatment year 2	.318	.928**

* Mean difference is significant at the 0.05 level.
** Mean difference is not significant at the 0.05 level.

between the groups. The post hoc showed that the treatment group at year three had significantly higher mean scores than year one ($p = .001$). The post hoc also revealed a significant difference between those who participated at 2 years of training and at 1 year of training ($p = .014$). This suggests that participating in professional development in the PESTL program for 2 or more years may increase the ability of elementary school teachers to align their instruction with the Utah state core standards.

The post hoc analysis did not reveal a significant difference with teachers who participated at 3 years of professional development training and at 2 years of professional development training within the PESTL program. The mean score was higher after 3 years ($M = 4.77$) when compared to 2 years ($M = 4.45$); however, this difference was not statistically significant ($p = .928$). This may be due to the small sample size of treatment teachers or perhaps the professional development was not effective in year three for developing the practice of content alignment. These findings support the hypothesis that teachers who receive two or more years of professional development are more likely to have enhanced science instructional practices in Content Alignment.

Addressing Science Misconceptions

The addressing misconceptions construct consists of one observational item rated on a 5-point scale, which is that science misconceptions were appropriately addressed by the teacher. Teachers were given a score for this indicator as outlined on the observation tool and noted by the evaluator. A full description of the PESTL Observation Protocol rating scale is included in Appendix B.

To address question two and specifically subquestion 2e, an independent *t* test was used to determine differences between the control group and treatment year one. The researcher chose to use *t* tests rather than ANOVAs because the control teachers were not the same as the treatment teachers, therefore the control scores did not qualify for repeated measures analyses. A Bonferroni correction was made for the *p*-value significance level in order to avoid a spurious positive and to account for the number of comparisons being performed. For the Bonferroni correction, the alpha values were lowered to $\alpha = .0125$.

In the independent *t* tests the independent categorical variable was group and the dependent continuous variable was the addressing misconceptions score. Addressing misconceptions scores from observations from year one of treatment were compared to the addressing misconceptions scores from the control group.

There was no significant difference between the control group and treatment year one $t(40) = -0.26$, $p = .797$. The treatment group had higher addressing misconceptions mean scores (*M* = 2.36) after 1 year compared to the control teachers (*M* = 2.25); however, these differences were not statistically significant ($p = .797$). This may be due to the small sample size of treatment teachers or may indicate the professional development year one was not effective in changing practice. This finding does indicate that teachers who participate in PESTL need two or more years of professional development to significantly increase their effective science instructional practices of addressing misconceptions. The results are that teachers who participate in PESTL professional development do not have significantly higher scores in addressing misconceptions after 1 year when compared to a control group of nonparticipating teachers.

A one-way repeated measures ANOVA was used to examine treatment group change over time in the practice of addressing misconceptions. The independent categorical variable was year of treatment and the dependent continuous variable was the addressing misconceptions score. Addressing misconceptions scores from observations of the treatment group from years one, two, and three were compared over time. Table 4.9 reports the finding of the repeated measures ANOVA.

The ANOVA revealed significant differences between the groups, $F(2, 63) = 11.508$, $p = .000$. Bonferroni post hoc analysis was used to examine differences between the groups. The post hoc showed that the treatment group at year three had significantly higher mean scores than at year one ($p = .000$), and at year two ($p = .044$). This suggests that participating in sustained professional development may increase the use of Addressing Misconceptions in the classroom science instruction of elementary school teachers.

Table 4.9

Repeated Measures ANOVA for Addressing Misconceptions of Treatment Group Over 3 Years

Group	Comparison year	Mean difference	*p*
Treatment year 1	Treatment year 2	-.909	.077**
	Treatment year 3	-1.909*	.000*
Treatment year 2	Treatment year 1	.909	.077**
	Treatment year 3	-1.000*	.044*
Treatment year 3	Treatment year 1	1.909*	.000*
	Treatment year 2	1.000*	.044*

* Mean difference is significant at the 0.05 level.
** Mean difference is not significant at the 0.05 level.

The Bonferroni revealed no significant difference between those who participated at year two of professional development training and at year one of professional development training (p = .077). The mean score was higher at year two (M = 3.27) when compared to year one (M = 2.36); however, these differences were not statistically significant. Among other factors, this may be due to the difficulty many elementary teachers have in practicing the component of addressing misconceptions in classroom instruction.

These results suggest that participating in sustained professional development does increase the practice of addressing misconceptions in the classroom science instruction of elementary school teachers. This finding supports the hypothesis that teachers who receive two or more years of professional development in the PESTL program are more likely to have enhanced science instructional practices in addressing misconceptions.

The findings suggest that participating in sustained professional development does increase the practices of the specific components of reformed science education in the classroom science instruction of elementary school teachers. These finding support the hypothesis that teachers who receive 3 years of professional development are more likely to have enhanced science instructional practices. It appears there are significant differences between groups. The treatment teacher scores after 3 years of professional development were significantly higher than treatment teacher scores after 1 year of professional development in all five components of reformed science education.

Research Question Three

Research question three asked, "How does sustained professional development of 3 years transform overall classroom science instruction and the specific components of reformed science instruction of elementary school teachers over time?" The researcher analyzed the data gathered from both the Summary of Effective Science Instruction and PESTL Observation Protocol to answer this question. Four subquestions for question three were identified as follows.

3a. What components of reformed science instruction are most influenced by sustained professional development over time?

3b. What components of reformed science instruction are least influenced by sustained professional development over time?

3c. Are there patterns of change among the components of reformed science instruction of elementary school teachers over time?

3d. Are there interaction between performance in overall classroom science instruction and the practice of specific components of reformed science instruction of elementary school teachers over time?

Independent *t* tests were used to compare the control group scores to the year one and year three treatment scores. The researcher chose to use *t* tests rather than ANOVAs because the control teachers were not the same as the treatment teachers; therefore, the control scores did not qualify for repeated measures analyses. A Bonferroni correction was made for the *p*-value significance level in order to avoid a spurious positive and to account for the number of comparisons being performed. For the Bonferroni correction, the alpha values were lowered to $\alpha = .0125$. Treatment year one scores were compared to the control scores to examine patterns of change over time. Treatment year three scores were compared to control scores to examine most and least influenced components of reformed science education.

One-way repeated measures ANOVAs with Bonferroni post-hoc were used for examining observation scores within the treatment group over time. These post-hoc analyses were used to examine the patterns of change among PESTL components and interactions between overall instruction and PESTL components.

Most Influenced Components of Reformed Science Instruction

The goal was to investigate the most influenced PESTL components. To address question three and specifically subquestion 3a, independent *t* tests were used to determine differences between the control group and treatment group year three. The control group scores were used to establish nontreatment compared to year three scores, which established the end of treatment scores.

In the independent *t* tests the independent categorical variable was group and the dependent continuous variables were the PESTL components. PESTL component scores from observations from year three of treatment were compared to the PESTL component scores from the control group. The researcher chose to use *t* tests rather than ANOVAs because the control teachers were not the same as the treatment teachers, therefore the control scores did not qualify for repeated measures analyses.

To determine what components of reformed science instruction were most influenced by sustained professional development over time, mean scores, mean differences, percentage of increase in mean scores from the five components of the PESTL Observation Protocol were reported. Scores from the control group and treatment year three were included. Table 4.10 displays the most and least influenced components of reformed science instruction.

Independent *t* tests were conducted on each of the five components separately. The independent *t* tests showed that the treatment group year three had significantly higher mean scores than the control group in all five components: Talk and argument $t(40) = -8.85$, $p = .000$, Investigation $t(40) = -8.73$, $p = .000$, Modeling $t(40) = -5.92$, $p = .000$, Content Alignment $t(40) = -3.62$, $p = .001$, and Addressing Misconceptions $t(40) = -4.52$, $p = .000$. Additionally, the mean differences illustrated greater gains in some components when compared to others. The most influenced components displayed

Table 4.10

Most and Least Influenced Mean Scores Control Group to Treatment Year 3

Dependent variable	Control group *M*	Treatment year 3 *M*	Mean difference	Increase in mean
Talk and argument	11.40	23.05	11.65	47%
Investigation	4.65	11.95	7.30	49%
Modeling	4.00	10.36	6.36	42%
Content alignment	3.15	4.77	1.62	32%
Addressing misconceptions	2.25	4.27	2.02	40%

higher percentages of difference in mean scores between control and year three of treatment.

The component of reformed science instruction that was most influenced by sustained professional development over time was Investigation, which displayed a 49% higher difference in mean scores between the control scores and year three of treatment scores. The second most influence component was talk and argument, which displayed a 47% higher difference in mean scores between the control scores and year three of treatment scores. The third most influence component was modeling, which displayed a 42% higher difference in mean scores between the control scores and year three of treatment scores. The fourth most influence component was addressing misconceptions, which displayed a 40% higher difference in mean scores between the control scores and year three of treatment scores.

The PESTL program's focus on the components of investigation, talk and argument, and modeling during the summer sessions appears to have impacted the practice of these components in the classroom. These findings further support the hypothesis that teachers who receive sustained professional development are more likely to have enhanced science instructional practices and that variance in most influenced components of PESTL exist. It appears there are significant differences between groups. The treatment teacher scores after 3 years of professional development were significantly higher than the control group scores.

Least Influenced Components of Reformed Science Instruction

The goal was to investigate the least influenced PESTL components. To address question three and specifically subquestion 3b, independent *t* tests were used to determine differences between the control group and treatment group year three.

In the independent *t* tests the independent categorical variable was group and the dependent continuous variables were the PESTL components. PESTL component scores from observations from year three of treatment were compared to the PESTL component scores from the control group. The researcher chose to use *t* tests rather than ANOVAs because the control teachers were not the same as the treatment teachers; therefore, the control scores did not qualify for repeated measures analyses.

To determine what components of reformed science instruction are least influenced by sustained professional development over time, mean scores, mean differences, percentage of increase in mean scores from the five components of the PESTL Observation Protocol were reported (see Table 4.10). Scores from the control group and treatment year three were included.

Independent *t* tests were conducted on each of the five components separately. The independent *t* test showed that the treatment group year three had significantly higher mean scores than the control group in all five components: Talk and argument $t(40) = -8.85$, $p = .000$, Investigation $t(40) = -8.73$, $p = .000$, Modeling $t(40) = -5.92$, $p = .000$, Content Alignment $t(40) = -3.62$, $p = .001$, and Addressing Misconceptions $t(40) = -4.52$, $p = .000$. Additionally, the mean differences illustrated smaller gains in one component when compared to the others. The least influenced component displayed a lower percentage of difference in mean scores between control and year three of treatment. The component of reformed science instruction that was least influenced by sustained professional development over time was content alignment, which displayed only a 32% higher difference in mean scores between the control scores and year three of treatment scores.

These findings suggest that teachers who receive sustained professional development are more likely to have enhanced science instructional practices and that variance in least influenced components of PESTL exist. The component of content alignment was found to be the least influenced by sustained professional development. This may be due to the emphasis already placed on covering the state core in elementary teaching practices for all teachers.

Patterns of Change Among the Components of Reformed Science Instruction

The goal was to investigate patterns of change among PESTL components. To address subquestion 3c, independent *t* tests were used to determine differences between the control group and treatment group year one. The researcher chose to use *t* tests rather than ANOVAs because the control teachers were not the same as the treatment teachers, therefore the control scores did not qualify for repeated measures analyses. A Bonferroni correction was made for the *p*-value significance level in order to avoid a spurious positive and to account for the number of comparisons being performed. For the Bonferroni correction, the alpha values were lowered to $\alpha = .0125$.

Independent *t* tests were conducted on each of the five components separately. The independent *t* test showed that the treatment group year one did not have significantly higher mean scores than the control group in any of the five components: talk and argument $t(40) = -.631$, $p = .531$, investigation $t(40) = -1.66$, $p = .106$, modeling $t(40) = -2.26$, $p = .030$, content alignment $t(40) = -0.95$, $p = .346$, and addressing misconceptions $t(40) = -0.26$, $p = .797$.

One-way repeated measures ANOVAs with Bonferroni post-hoc were used to determine differences within the treatment group over time.

To examine patterns of change, independent *t* tests were conducted using group as the independent categorical variable and the PESTL components scores as the dependent continuous variables. PESTL component scores from observations from year one of treatment were compared to the PESTL component scores from the control group.

In the one-way repeated measures ANOVAs, the independent categorical variable was year of treatment and the dependent continuous variables were the PESTL components. PESTL component scores from observations of the treatment group from year one, year two, and year three were compared over time.

To examine patterns of change among the components of reformed science instruction of elementary school teachers over the 3-year term of the professional development program, significant differences among the five components of the PESTL Observation Protocol were charted in frequency tables (see Table 4.11). Scores from the control group and treatment group year one, year two, and year three were included. Patterns of change in each PESTL component from nontreatment to year one, year one to year two, and year one to year three were analyzed and described.

The frequency table illustrates patterns of change over time and reveals how PESTL sustained professional development of 3 years transformed overall instruction and specific components of instruction over time. Significant improvements were made in Modeling after 1 year of treatment when compared to nontreatment, however with the

Table 4.11

Patterns of Change in Components of Reformed Instruction

Group	Compared	Talk and argument *p*	Investigation *p*	Modeling *p*	Content alignment *p*	Addressing misconceptions *p*
Year 1	Control	.531**	.106**	.030**	.346**	.797**
	Year 2	.000*	.007*	.145**	.014*	.077**
	Year 3	.000*	.000*	.002*	.001*	.000*
Year 2	Year 1	.000*	.007*	.145**	.014*	.077**
	Year 3	.002*	.061**	.392**	.928**	.044*
Year 3	Year 1	.000*	.000*	.002*	.001*	.000*
	Year 2	.002*	.061**	.392**	.928**	.044*

* The mean difference is significant at the 0.05 level for ANOVA Bonferroni post-hoc and .0125 for *t* test Bonferroni adjustment.

** The mean difference is not significant.

Bonferroni correction the difference was reduced to nonsignificant. Significant improvements were made in Talk and argument, Investigation, and Content alignment between year one and year two of treatment; however, significant improvements were made in all five components of reformed instruction concomitantly only after 3 years of treatment when examining the treatment group over time. These findings further support the hypothesis that teachers may benefit from sustained professional development of 3 years in order improve in all components of reform science education.

Interaction Between Overall Science Instruction and Components of Reformed Science Instruction

The goal was to investigate interactions and linear relationships between overall instruction and PESTL components. To address question 3d, descriptive statistics were used to determine mean scores of the control group and treatment group year one, year two and year three. Each construct was reduced to an equalized mean out of five through simple division in order to analyze interactions and relationships between the individual PESTL components and the overall measures of the Summary Judgment and the PESTL Total.

Mean scores were equalized to the common denominator of five total points so to compare numbers and trends equally. In order to equalize the mean scores, each component total mean score was divided by the number of indicators included in the component. The PESTL Total included 13 five-point indicators so the mean score was divided by 13. Talk and argument included five 5-point indicators, so the mean score was divided by five. Investigation included three 5-point indicators, so the mean score was divided by three. Modeling included three 5-point indicators, so the mean score was divided by three. Summary, content alignment, and addressing misconceptions each had one 5-point indicator, so each maintained its original mean score. Mean scores from the control group year one and treatment group year one, year two, and year three were included (see Table 14.12).

To determine interaction between performance in overall classroom science instruction and the practices of the PESTL components of elementary school teachers over time, mean scores from the five components of the reformed science were plotted on graphs. These mean scores were compared to Summary Judgment of Instruction and PESTL Observation Total.

The graph illustrates interactions over the 3-year time period of professional development. Interactions between performance in overall classroom science instruction and each PESTL component from control to year one, year one to year two, and year two

Table 4.12

Equalized Mean Scores Control Group to Treatment Year One, Two, and Three

Dependent variable	Control *M*	Year 1 *M*	Year 2 *M*	Year 3 *M*
Summary	2.20	2.73	3.41	4.14
PESTL total	1.96	2.51	3.62	4.42
Talk and argument	2.28	2.48	3.78	4.61
Investigation	1.55	2.21	3.23	3.98
Modeling	1.33	2.23	2.92	3.45
Content alignment	3.15	3.55	4.45	4.77
Addressing misconceptions	2.25	2.36	3.27	4.27

Note. All mean scores have been equalized to the common denominator of five. This was done by dividing the variable mean by the number of indicators within that variable. Each indicator was measured on a 5-point scale.

to year three were analyzed and described. Figure 4.1 and Figure 4.2 illustrate trends, linear relationships, and interactions between the components of reformed science instruction and the measures of overall science instruction, which includes the Summary Judgment of Instruction and PESTL Total.

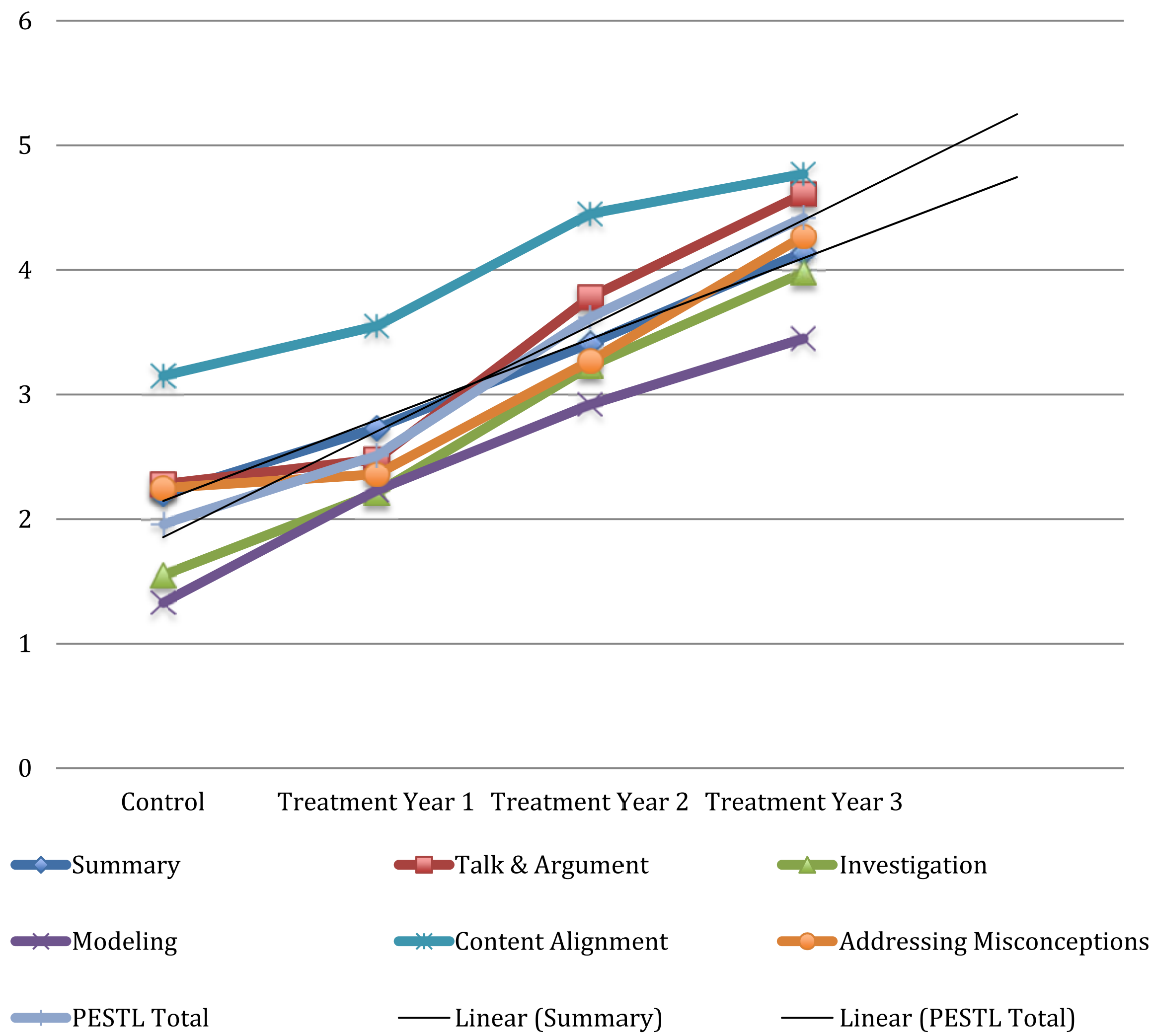

Figure 4.1. Interactions between summary judgment of instruction, PESTL total, and components of reformed instruction with equalized mean scores.

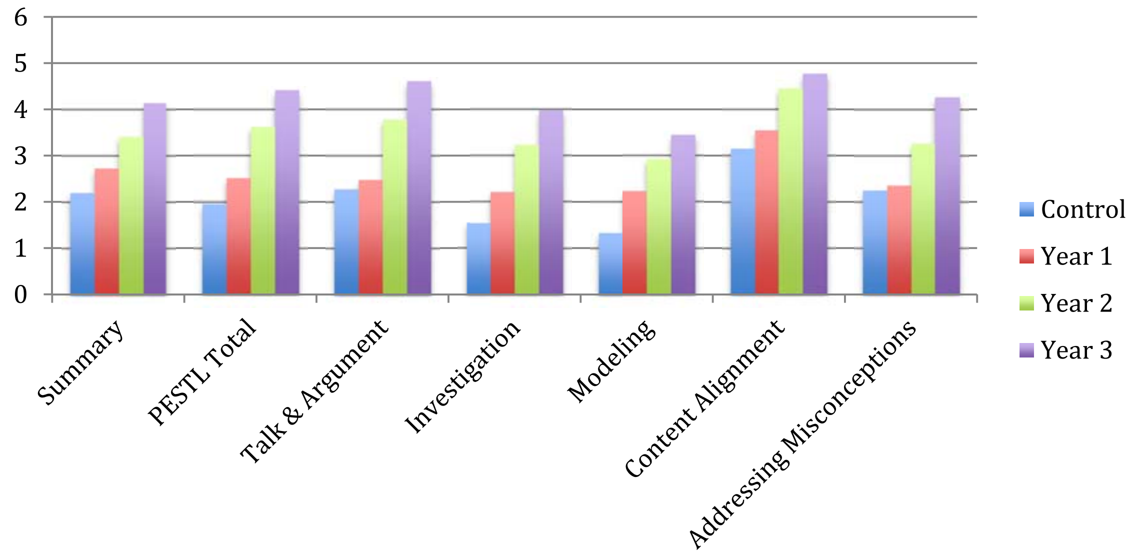

Figure 4.2. Trends of summary judgment of instruction, PESTL total, and components of reformed instruction with equalized mean scores.

Figures 4.1 and 4.2 illustrate a positive linear relationship between the summary judgment of instruction, the PESTL total, and the specific components of reformed science instruction scores of elementary school teachers over time. This finding suggests that trends are present. The gradual positive trend illustrated for the summary scores is similar to that of content alignment and addressing misconceptions, whereas modeling and investigation scores illustrate greater positive direction than the summary scores. These finding suggests that summary scores are less influence by sustained professional development, as content alignment and addressing misconceptions were the least influence components, and modeling and investigation were most influenced (see Table 4.10).

The steep treatment trend of the PESTL total is most similar to the talk and argument trend, both measures illustrate a step-like positive trend showing increasing gains during year two and three of treatment. This finding suggests that there is a strong interaction between the PESTL total and talk and argument trends.

These findings further support the hypothesis that teachers who participated in PESTL sustained professional development likely have improved practices in overall science instruction and specific components of reformed science education and that interactions between overall instruction and PESTL components exist.

Summary

The purpose of this study was to determine the extent to which sustained professional development in elementary science education affects classroom instruction over a 3-year time period. The data analyses utilized were independent *t* tests and one-way ANOVAs. Independent *t* tests were used for examining data between the control group and treatment group. One-way repeated measures ANOVAs with Bonferroni post-hoc were used for examining change over time within the treatment group.

The findings illustrated that teachers, who participated in 1 year of PESTL professional development, did not have significantly higher scores in either the Summary Judgment of Science Instruction or the PESTL total when compared to the control teachers. When looking at changes over time within the treatment group, participants in year two had significantly higher scores in PESTL Observation Protocol total when compared to year one of treatment. Participants in year three of professional development had significantly higher scores in the Summary Judgment of Science Instruction and PESTL Observation Protocol total when compared to 1 year and 2 years of treatment.

In specific components of reformed science instruction findings illustrated that teachers, who participated in 1 year of PESTL professional development, does not have significantly higher scores when compared to nontreatment, except in the component of modeling. When looking at changes over time within the treatment group, participants in year two has significantly higher scores in talk and argument, investigation, and content alignment when compared to year one of treatment. Participants in year three of professional development had significantly higher scores in the talk and argument, and addressing misconceptions when compared to 2 years of treatment. Finally, participants in 3 years of professional development had significantly higher scores in all five components when compared to 1 year.

Findings further illustrated that there were most and least influenced components of PESTL, patterns of change among PESTL components, and that interactions between overall instruction scores and PESTL components did exist. When comparing the mean scores between the control group and year three of treatment, the component that was most influence by 3 years of sustained professional development was investigation, next was talk and argument, followed by modeling, and then addressing misconceptions. The least influenced component was content alignment.

When scores from the control group and treatment group year one, year two, and year three in all seven measures were entered into a frequency table, patterns of change were revealed. The findings illustrated that teachers, who participated in 1 year of PESTL professional development, did not have significantly higher scores when compared to nontreatment, except in the component of modeling. Participants in year two showed significantly higher scores in PESTL Observation Protocol total, talk and argument, investigation, and content alignment when compared to year one of treatment. Participants in 3 years of professional development had significantly higher scores in Summary Judgment of Science Instruction, PESTL Observation Protocol total, talk and argument, and addressing misconceptions when compared to 2 years of treatment. Finally, participants in 3 years of professional development had significantly higher scores in all seven measures when compared to 1 year.

Interactions between overall instruction scores and PESTL components were revealed. A positive linear relationship between the summary judgment of instruction, the PESTL total and the specific components of reformed science instruction scores of elementary school teachers over time. The gradual positive trend illustrated for the summary judgment of instruction is similar to that of content alignment and addressing misconceptions, whereas the trend of modeling and investigation illustrate greater positive direction. The steep treatment trend of the PESTL Total is most similar to the Talk and argument trend both illustrate a step-like positive trend showing increasing gains during year two and three of treatment.

These findings support the hypothesis that sustained teacher professional development, designed to meet recommendations for reform in elementary science education, can transform classroom instruction over a 3-year time period.

CHAPTER 5

DISCUSSION

The researcher completed the data analyses and reviewed the results of the study. The analyses revealed significant improvements in classroom science instruction scores for the teachers who participated in all 3 years of the PESTL program. The researcher found that none of the components of instruction improved significantly after 1 year of treatment, all measures took 2 or even 3 years to demonstrate significant improvement. The findings suggested that 3 years of PESTL professional development was ideal for significantly improving overall science instruction and all five components of reformed science instruction concomitantly. As the researcher reflected on the findings, she considered the implications of the study and the impact this study may have on future professional development programs as well as future research in elementary science teaching and learning.

Introduction

The purpose of this study was to determine whether sustained professional development would affect the classroom science instruction of elementary school teachers. The central question of this study was, does sustained teacher professional development, which was designed to meet recommendations for reform in elementary science education, transform classroom instruction over a 3-year time period? In regards to this question, three questions, with related subquestions, were explored.

1. Does sustained professional development of 3 year, which was designed to meet recommendations for reform in science education, transform overall classroom science instruction of elementary school teachers?

2. Does sustained professional development of 3 years transform the instructional practice of elementary school teachers in specific components of reformed science education?

2a. Does sustained professional development effect the practice of classroom discussion and argumentation in the science instruction of elementary school teachers?

2b. Does sustained professional development effect the practice of scientific investigation in the classroom science instruction of elementary school teachers?

2c. Does sustained professional development effect the practice of teacher and student modeling in the classroom science instruction of elementary school teachers?

2d. Does sustained professional development effect the practice of aligning science instruction to core standards in the classroom instruction of elementary school teachers?

2e. Does sustained professional development affect the practice of appropriately addressing student science misconceptions in the classroom instruction of elementary school teachers?

3. How does sustained professional development of 3 years transform overall classroom science instruction and the specific components of reformed science instruction of elementary school teachers over time?

3a. What components of reformed science instruction are most influenced by sustained professional development over time?

3b. What components of reformed science instruction are least influenced by sustained professional development over time?

3c. Are there patterns of change among the components of reformed science instruction of elementary school teachers over time?

3d. Are there interactions between performance in overall classroom science instruction and the practice of specific components of reformed science instruction of elementary school teachers over time?

This study examined the effects of sustained professional development on the classroom science instruction of elementary school teachers. The study connected the paradigm of educational reform, specifically in elementary science education, and used the transformative learning theory as a theoretical lens.

Data used for this study were preexisting. Data were collected during classroom observations of 22 cluster sampled treatment teachers who were evaluated each year over the 3 years of professional development. To establish a baseline for nontreatment, 20 convenience-sampled control teachers were observed once during the first year of the study. PESTL evaluators made 45-minute classroom observations of science instruction during the regular school day using two instruments: the PESTL Observation Protocol on Instructional Effectiveness and the Summary Judgment of Science Instruction. Data were collected over a 3-year time period from Winter 2008 to Spring 2011.

This study contributes to the literature in many ways. First, it looked at professional development in science education as a transformational learning process with social constructivist considerations in reforming instruction. This theoretical lens enabled the researcher to examine and compare how the different characteristics of a 3-year sustained professional development program transformed participating teachers' instruction over time. Synthesis studies reported that reform in science teaching is best accomplished by viewing teacher professional development through a transformative lens (NRC, 2003; Stein et al., 1999). Most previous research on elementary science professional development have examined groups of teachers within a short time, a single year or less, preventing the ability to study participants' transformation of instruction over time.

The PESTL professional development design is transformational in nature. Participants are given disorienting dilemmas in relation to their understanding of science concepts and science teaching practice. Following the disorienting dilemma the participants reexamine their beliefs about the content and teaching practice they use in the classroom. Participants then discuss and reflect on the new learning they have acquired. Because the program is sustained, the participants have time to change their instruction to be more inclusive of effective science teaching and learning practice.

Transformational learning consists of two basic kinds of learning modes: instrumental and communicative (Mezirow & Taylor, 2009). Instrumental learning focuses on learning through task-oriented problem solving and cause-effect relationships. In the instrumental mode understandings are validated by empirical evidence to ascertain the truth of a belief, association, concept, value, feeling, or world-view (Mezirow & Taylor, 2009). Communicative learning focuses on how individuals communicate their feelings, needs, and desires (Taylor, 2000). In the communicative mode a contested belief, association, concept, value, feeling, or world view is validated or justified through dialogic discourse. This discourse facilitates social constructivist process in which adults collaborate to develop understanding and meaning. Both instrumental and communicative modes of learning relate to the findings of this study. The transformational learning theory is revisited in this chapter to reveal how it helps explain and inform the findings of this study.

Second, this study contributed to the literature by examining recent reformed instructional models in science education and instruments used to measure effectiveness in reformed science instruction. The NSES outlined a new vision of elementary science education through promoting changing emphases in science education. These initiatives and standards promote science literacy and educational reform through teacher professional development. In a synthesis of findings report, *Taking Science to School*, the NRC (2007) emphasized five key teaching models and practices observable and measurable in effective reformed elementary science instruction: (a) talk and argument, (b) modeling and representations, (c) investigations and inquiry, (d) alignment to science core concepts, and (e) appropriately addressing science misconceptions.

Many previous studies on elementary science professional development have measured instruction with outdated or general instruments that do not measure the practices of reformed science instruction. The PESTL Observation Protocol is a classroom observation tool designed specifically to evaluate the five components of reformed science instruction in an elementary setting. This tool was used in this study to gather data on the transformation of reformed instructional practices of teachers participating in sustained professional development. A second more established tool, the Summary Judgment of Science Instruction, which was adapted from HRI's Capsule

Rating of the Quality of Instruction (HRI, 2001), was used to evaluate overall science instruction. The researcher used *t* tests to examine differences between a control group and treatment group, whereas repeated measure ANOVAs were used to examine change over 3 years of professional development within the treatment group.

Finally, this study investigated professional development characteristics that have proven effective in transforming the classroom science instruction of elementary school teachers. The NSES' changing emphasis in science teaching and standards for professional development suggest the need for professional development to be long-term, focused on building teacher science literacy, rooted in research on how children learn science, and embedded in teaching practice in the classroom (NRC, 1996).

The summary of findings, recommendations for practice, limitations, and implications for future research based on the findings of this study are discussed in this section.

Summary of Findings

This section summarizes the findings presented in this research study. Due to the multiple research questions included in this study, this summary will be organized into three sections: (a) transformation of overall classroom science instruction; (b) transformation of instructional practice in specific components of reformed science education; (c) comparisons, patterns and interactions of instructional practices over time.

Transformation of Overall Classroom Science Instruction

This study builds on a research base investigating elementary science teacher professional development programs. Current studies on the effects of teacher development on classroom instruction report mixed findings on the effectiveness of teacher development on changing overall elementary science teaching practice. Many previous studies on this topic have looked at short-term, a few weeks to a year of professional development (Drits & Stark, 2011; Minuskin, 2009; Santau, 2008). Few studies have looked at the same group of teachers for 3 years of sustained professional development.

To determine the effects of sustained professional development on the overall science instructional practices of elementary teachers in this study, descriptive statistics and repeated measures analysis of variance were performed on two data sets. The first set was from the Summary of Science Instruction ratings and the second set was from the PESTL Observation Protocol Total scores. Statistical analysis of the data revealed that there were no significant differences in overall science instruction measured in Summary of Science Instruction or PESTL Total scores between the teachers who participated in 1 year of PESTL when compared to the control group.

When examining the participants over time there were significant differences. With the Summary Judgment ratings the treatment group year three had significantly higher mean scores than year one (p = .000) and year two (p = .044). However, there were no significant difference between those who participated in 1 year of professional development training and those who participated in 2 years of professional development training (p = .066). This suggests that participating in at least 2 years of sustained professional development is ideal for increasing the overall effectiveness of the classroom science instruction of elementary school teachers.

With the PESTL Total Scores the treatment group year three had significantly higher mean scores than year one (p = .000), and year two (p = .003). Likewise, there were significant differences between those who participated in 2 years of professional development training and those who participated in 1 year (p = .000). These findings support the hypothesis that teachers who receive sustained professional development will have improved effectiveness in overall science instruction.

These findings are consistent with current literature on sustained science professional development. Among the more rigorous studies of professional development for teachers of science are those of a longitudinal study of sustained professional development by the Merck Institute for Science Education (Corcoran et al., 2003); the NSF-funded studies of systemic reform in mathematics and science (Supovitz & Turner, 2000; Weiss et al., 2003); and evaluations of the federal Eisenhower mathematics and science professional development program (Garet et al., 1999). These studies examined both elementary and secondary science education and found significant improvement over time with sustained professional development. Studies found that some gains are found after 2 years of professional development, but significant improvements are best achieved and maintained with 3 years (Corcoran et al., 2003; Supovitz & Turner, 2000). These studies explained that sustain professional development for science education allows teachers to integrate new knowledge and strategies into their practice and to reflect on their experiences (Corcoran et al., 2003; Garet et al., 1999; Supovitz & Turner, 2000; Weiss et al., 2003). This explanation also applies to this study.

Transformational learning theory posits that adult learning takes time (Mezirow, 2000). This theory offers a possible explanation that is consistent with the data examined in this study. Transformation learning theory also helps inform the findings of this study in relation to sustained professional development providing the time participants needed to practice more effective overall science instruction in the classroom. These findings further suggest that participants experienced both instrumental and communicative modes of transformational learning (Mezirow & Taylor, 2009).

Instrumental learning, which is validated by *experience and outcomes*, was presented in the PESTL Summer Seminars, Mid-winter Institutes, and Content Courses each year. Over time these professional development components provided instrumental learning experiences to participating teachers. PESTL participants likely relied on instrumental learning over time to reflect on transformations in their understanding of science concepts and science teaching practices in their classrooms.

Findings from this study suggest that as teachers experienced instrumental transformation through sustained professional development, their overall classroom science teaching practice changed. Current research in teacher development supports the extensive integration of professional development experiences with classroom instruction (Birman et al., 2000; Lemke, 2001; NRC, 1999, 2011; Stein et al., 1999). When this integration is effective, it allows participating teachers the opportunity to examine the basis of what is being learned in the context that it will be employed (Campbell, 2012).

PESTL is designed to develop science teaching and learning in the context in which it will be employed, the classroom, so participants can experience the learning and then discuss and reflect on the outcomes. This design aligns with transformational instrumental learning. A specific example of instrumental learning taking place in PESTL was revealed in the practice of Investigation in classroom instruction of participants. Findings suggest that as teachers experienced well designed investigations in the professional development components of PESTL they incorporated more indicators of effective investigations in their classroom instruction. Over time participants demonstrated a significant increase in effective Investigation practices including: instructing students to investigate problems, make observations, formulate testable questions, identify variables, record data, make inferences, classify, discuss limitations of findings, and analyze data.

Communicative learning, which focuses on *expression* of ideas and viewpoints through discourse, was provided in all five PESTL components. PESTL participants had many opportunities to express their thoughts and beliefs in the professional development activities. Innovative approaches to professional development encourage teachers to examine basic questions about what it means to be a teacher through connecting the professional development experience with instructional practice in the classroom (Campbell, 2012; NRC, 2011). PESTL is designed to develop science teaching and learning through discussion and collaboration. This design aligns with transformational communicative learning.

Findings from this study suggest that as teachers experienced communicative transformation through sustained professional development, their overall classroom science teaching practice changed. A specific example of communicative learning taking place through sustained professional development was the change in Talk and argument in classroom instruction of participants. Findings suggest that as teachers experienced communicative learning in the professional development components of PESTL they incorporated more indicators of effective Talk and argument in their classroom instruction. Over time participants demonstrated significant increase in effective Talk and argument practices including: the number of interactions between teacher and student, number of times teacher actively extended student thinking, number of examples and analogies in presentations, supporting relevant inter-student discussion, and the use of accurate science language in instruction.

Transformation of Specific Components of Reformed Science Education

The second research question sought to determine the effects of sustain professional development on the five components of reformed science education. Previous studies have looked at individual components of science instruction (Crawford, 2000; Duschl & Osborne, 2002; Michaels et al., 2008; Ornek, 2008; Wu & Krajcik, 2006), but few studies have looked at multiple components in the same professional development program.

To determine the effects of sustained professional development on the multiple components of reformed science instructional practices of elementary teachers in this study, descriptive statistics and repeated measures analysis of variance were performed on five data sets gathered from classroom observations with the PESTL Observation Protocol. The five components include: (a) talk and argument, (b) modeling, (c) investigations, (d) content alignment, and (e) addressing misconceptions. With the exception of one component, modeling (p = .030), statistical analysis of the data revealed that there were no significant differences in the specific components of reformed science instruction scores between the teachers that participated in 1 year of PESTL when compared to the control group of teachers that did not participate. However, with the Bonferroni correction the significance for modeling was changed to nonsignificant.

Statistical analysis of the data revealed that there were significant differences over time with teachers who participated in PESTL. There were significant differences between 1 year and 2 years of professional development training in three components: talk and argument (p = .000), investigation (p = .007), and content alignment (p = .014). There were significant differences between 2 years and 3 years of professional development training in two components: talk and argument (p = .002) and addressing misconceptions (p = .044). There were significant differences between 1 year and 3 years of professional development training in all five components: talk and argument (p = .000), investigation (p = .000), modeling (p = .002), content alignment (p = .001), and addressing misconceptions (p = .000). These findings support the hypothesis that teachers who receive sustained professional development will improve in specific components of reformed science instruction over time.

These findings can be compared and contrasted to current literature on science education and elementary science professional development. Since there are no past studies that have examined the multiple components of reformed science education covered in this study, findings on each component will be examined separately. The five components include: (a) talk and argument, (b) modeling, (c) investigations, (d) content alignment, and (e) addressing misconceptions.

Talk and argument. In the past three decades, the role of language in the science curriculum has become prominent in science education literature (Dawes, 2004; Gee, 1989; Lemke, 1990; Yore et al., 2003). Reports on reformed science education revealed that effective teaching practices include a high ratio of interactions between students and the teacher as well as frequent opportunity for students to discuss scientific ideas with their peers (NRC, 2007).

This study revealed significant improvement all 3 years of professional development in Talk and argument. The concept of scientific talk and argumentation was introduced early in the PESTL professional development program and built upon each year. This likely explains the continuous improvement in its use amongst participants over time. The talk and argument scores also resembled the linear growth of the PESTL total scores over time. This shows a high correlation between improvements in talk and argument and the overall PESTL Observation Protocol rating. This may also be explained by the higher number of indicators in the talk and argument component of the tool; there were five indicators with a possible score of 25 for talk and argument compared to three or less indicators in the other four components. Therefore, talk and argument scores are more likely to correlate with the PESTL total scores.

In comparison to the other components of reformed science education talk and argument was the second most influenced component of the five components measured when compared to the control group. Patterns of change revealed that talk and argument in the treatment group had a higher mean score year one when compared to the control but it was not statistically significant. Scores were significantly higher year two and three when compared to year one and year three when compared to year two. Talk and argument showed a positive linear growth similar to the PESTL total scores and increased at a greater degree than the Summary Judgment scores over time.

Modeling. In the past two decades, science education experts have increasingly recognized the value of modeling and representations in the science education reform movement (AAAS, 1993; Giere, 1991; Gobert & Buckley, 2000; NRC, 1996, 2007). At present, models and modeling are considered integral parts of scientific literacy (S. Gilbert, 1991; J. Gilbert, 1993; Gilbert & Boulter, 1998; Linn & Muilenberg, 1996; Perkins, 1986). Studies on reformed science education revealed that effective teaching practices include making students' thinking visible through modeling and representations (Driel & Verloop, 1999; J. Gilbert, 1995; Gilbert et al., 2000; NRC, 2007).

This study revealed significant improvement in year one of treatment when compared to the control group and between year one and year three when looking at PESTL participants over time. The concept of scientific modeling was introduced the first year of PESTL professional development, but was not focused on until the third year. In contrast to the other components, which showed no significant difference between the control and year one of treatment, modeling did show significant differences. This may be due to the lack of background knowledge that most elementary teachers have with the concept of modeling in science. However, with the Bonferroni correction the significance for modeling was changed to nonsignificant.

The modeling rating for the control teachers was the lowest mean score of all seven constructs. The professional development in year one may have provided enough exposure to the concept of modeling to explain the significant differences between groups in year one. However, significant improvement over time was only found between year one and three within the treatment group. This illustrates the need for sustained professional development, especially for the component of modeling, as it appears to be a difficult practice of reformed science education to develop in elementary teachers.

In comparison to the other components of reformed science education modeling was the third most influenced of the five components measured when compared to the control group. Patterns of change revealed that modeling consistently had the lowest equalized mean scores with the control group and the treatment group over time. Modeling showed a positive linear growth similar to the Summary Judgment of Science Instruction scores, but not as steep as the PESTL Total scores over time.

Investigations and inquiry. Recent studies suggest that effective science instruction includes learning science through investigations and inquiry-based instruction (NRC, 2007). Most past research on professional development in elementary science teaching has focused on practice of inquiry-based instruction (Crawford, 2000; Drits & Stark, 2011; Keys & Bryan, 2000; Minuskin, 2009; Santau, 2008; Wu & Krajcik, 2006). Researchers who study inquiry-based instruction are not in favor of a predetermined procedure for teaching; however, they do agree that a systematic study of inquiry methods is necessary in promoting its use in science education reform (Crawford, 2000; Wu & Krajcik, 2006). The function of inquiry-based instruction is for students to find answers to questions by way of gathering data or evidence (Richardson & Liang, 2008). Studies on reformed science education and inquiry-based teaching revealed that effective teaching practices include making students' thinking visible through investigations (Michaels et al., 2008; NRC, 2007; Richardson & Liang, 2008).

This study revealed significant improvement in Investigation scores between year one when compared to years two and three of treatment. However, though mean scores were higher they were not statistically significant between the control group and year one or between year two and three of treatment. The use of investigation was discussed the first year in the PESTL professional development program, but was focused on most during year two. This may explain the higher gains between year one and two. This may also explain the slight leveling out of the gains between years two and three, as the focus of PESTL turned to modeling in year three.

In comparison to the other components of reformed science education Investigation was the most influenced component of the five components measured when compared to the control group. Patterns of change revealed that Investigation in the treatment group had a higher mean score year one when compared to the control but it was not statistically significant. Scores were significantly higher year two and three when compared to year one. There was no significant difference between year three when compared to year two. Investigation showed a positive linear growth similar to the PESTL total scores, and increased at a greater degree than the Summary Judgment scores over time.

Alignment to science core concepts. In the standards-based reform movement, science core curriculum drives classroom instruction. In the PESTL program the science core was based on the Utah Elementary Science Core Curriculum (USOE, 2002). The Utah Elementary Science Core Curriculum was designed using the AAAS's Project 2061: Benchmarks for Science Literacy (1993) and the National Academy of Science's NSES (1996) as guides to determine appropriate content and skills. The NRC recommends that effective science instruction align with science core concepts and standards (2007).

This study revealed significant improvement in content alignment scores between year one when compared to years two and three of treatment. However, though mean scores were higher they were not statistically significant between the control group and year one or between year two and three of treatment. Alignment to core concepts of science was discussed throughout the PESTL professional development program. With the current standards-based movement, much focus has been placed on teaching core content, so this practice is more customary to elementary teachers. This may explain the slightly flat gains over time with content alignment scores.

In comparison to the other components of reformed science education content alignment was the least influenced component of the five components measured when compared to the control group. Again, this may be due to the common use of this practice in all teaching. Patterns of change revealed that content alignment in the treatment group had a higher mean score year one when compared to the control but it was no statistically significant. Scores were significantly higher year two and three when compared to year one. There was no significant difference between year three when compared to year two. Content alignment showed a positive linear growth but it was much less steep when compared to the PESTL Total scores as well as the summary judgment scores over time. In the future PESTL may want to examine crosscutting concepts and reformed practices of teaching science concepts to better measure changes in this component.

Addressing science misconceptions. The past 20 to 30 years of research have shown that children come to school with a great capacity for learning in general and learning science in particular (Metz, 1995; NRC, 1999, 2007). Children typically have significant gaps in their understanding (as do many adults), and their unschooled reasoning abilities may lead them to draw erroneous conclusions (Michaels et al., 2008). Effective science instruction requires the teacher to appropriately address science misconceptions when they occur.

This study revealed significant improvement in addressing misconceptions scores between years one and two when compared to year three of treatment. However, though mean scores were higher they were not statistically significant between the control group and year one or between years one and two of treatment. Addressing misconceptions of science was discussed throughout the PESTL professional development program; however, this component did not show significant improvement until after 2 years of professional development. It appears to be difficult to develop this practice in elementary teachers. This further illustrates the need for sustained professional development of two or more years.

In comparison to the other components of reformed science education addressing misconceptions was the second least influenced component of the five components measured when compared to the control group. This may be due to the time it takes for teachers to develop the scientific literacy and confidence to address science misconceptions when they arise during instruction. Patterns of change revealed that addressing misconceptions in the treatment group had a higher mean score year one when compared to the control but it was no statistically significant. Scores were significantly higher year two and three when compared to year one. There was no significant difference between year two when compared to year one.

Addressing misconceptions showed very little growth the first year; however, a steep positive linear growth was illustrated over year two and three of treatment. Gains in year two and three were much higher in addressing misconception in comparison to the PESTL Total scores as well as the summary judgment scores. These findings illustrate the effectiveness of sustained professional development in transforming the practice of appropriately addressing science misconceptions in the instruction of elementary teachers.

Transformational learning theory and question two findings. In relation to the transformational learning theory, findings for question two suggest that in order to transform the specific components of science education, participants required instrumental modes of learning to change some components and communicative modes of learning to change others (Mezirow & Taylor, 2009). Instrumental learning likely led to transformation in the components of investigation, modeling, and content alignment.

These three components rely heavily on experience and outcomes, which are authenticated through instrumental learning.

Communicative learning likely led to transformation in the components of talk and argument and addressing misconceptions in classroom instruction. These two components rely heavily on expression, which are supported by communicative learning. Over time the PESTL professional development program provided communicative learning experiences in talk and argument and addressing misconceptions to participating teachers, who in turn changed their classroom instruction to more effectively include these practices. Findings suggest that as these teachers experienced this communicative transformation, their practice of utilizing the specific components of reformed science education changed.

Comparisons, Patterns and Interactions of Instructional Practices over Time

The final research question in this study examined comparisons, patterns and interactions between overall science instruction and the specific components of instruction over time. The purpose of this analysis was to determine how science instruction changes over time. No previous studies could be identified that examined the relationships between the components of reformed science or the interactions between overall instruction and specific components.

To determine how sustained professional development of 3 years transformed overall classroom science instruction and the specific components of reformed science instruction of elementary school teachers over time, descriptive statistics were performed on seven data sets gathered from classroom observations with the Summary Judgment of Instruction and PESTL Observation Protocol. Statistical analysis of the data revealed how instruction changed over time in regard to which components of reformed science instruction were most and least influenced by sustained professional development with teachers participating in PESTL.

Most influenced components. The component of reformed science instruction that was most influenced by sustained professional development over time was Investigation, which displayed a 49% higher difference in mean scores between the control scores and year three treatment scores. The second most influenced component was talk and argument, which displayed a 47% higher difference in mean scores between the control scores and year three treatment scores. The third most influenced component was modeling, which displayed a 42% higher difference in mean scores between the control scores and year three treatment scores (see Table 4.10).

The PESTL program developed understanding and practices of all five components of reformed science education throughout the program; however, each year the summer workshops focused on one specific component. During year one, the practice of talk and argumentation in science teaching was focused on. During year two, the practice of investigation and inquiry-based instruction was focused on. During year three, the practice of modeling and representations in science teaching was focused on. This may explain why these components were the three most influenced over the 3 years of PESTL professional development.

Least influenced components. The least influenced component was content alignment, which displayed a 32% higher difference in mean scores between the control scores and year three of treatment scores. The second least influenced component was addressing misconceptions, which displayed a 40% higher difference in mean scores between the control scores and year three treatment scores (see Table 4.10). These two components showed very different changes in mean scores over time. The mean scores for content alignment were the highest each time measured, whereas the mean scores for addressing misconceptions started in the middle and remained flat until year two of the program.

As mentioned above, the PESTL program developed understanding and practices of all five components of reformed science education throughout the program; however, the three most influence components were focused on specifically in the summer workshops; whereas, these two least influence components were not. Sustained professional development likely had less influence on content alignment because this practice is already more customary to elementary classroom instruction. As mentioned previously, the current standards-based movement places much focus on teaching core content. This may explain the higher mean scores in this component with the control group and each year of treatment. Since this practice of teaching was already established before treatment, it showed fewer gains and therefore was least influenced.

On the other hand, addressing science misconceptions is not customary in elementary classroom instruction. The addressing misconceptions component showed little improvement until after 2 years of professional development. This practice depends on the science literacy of the teacher, which takes time to develop. These findings suggest that the practice of addressing misconceptions is more difficult to develop in elementary teachers. Sustained professional development of at least 2 years is necessary to improve this component. These findings support the hypothesis that teachers who receive sustained professional development will demonstrate changes in instruction over time and will have comparisons between least and most influenced components of reformed science instruction.

Patterns of change. Statistical analysis of the data revealed how instruction changes over time in regard to patterns of change among the components of reformed science instruction of elementary school teachers. After 1 year of treatment, significant improvements were made in only one component, modeling, when compared to nontreatment. However, with the Bonferroni correction the significance was changed to non-significance. Significant improvements were made in talk and argument, investigation, and content alignment after 2 years of treatment when compared to 1 year. Significant improvements were made in talk and argument and addressing misconceptions after 3 years of treatment when compared to 2 years. Furthermore, significant improvements were made in talk and argument, investigation, modeling, content alignment, and addressing misconceptions after 3 years of treatment when compared to 1 year.

These patterns of change illustrated that little improvement is made in specific components of reformed science instruction with only 1 year of professional development. Modeling was the only component that demonstrated significant improvement. However, with the Bonferroni correction the significance was changed to non-significance. This difference is likely due to the fact that modeling is unfamiliar to most elementary school teachers; therefore the scores were extremely low for the control teachers in this component. Patterns further illustrated that 2 years of professional development made notable gains in scores, with three components demonstrating significant improvements. Modeling and addressing misconceptions were the two components that did not demonstrated significant improvement from year one to year two. These patterns further illustrate the difficulty in developing the practice of these two components (see Table 4.11).

Patterns revealed significant improvements in two components from year two to year three. The first was talk and argument, which had significant improvements all 3 years of treatment. The second was addressing misconceptions, which did not demonstrate significant improvement until year three. This revealed the need for 3 years of sustained professional development in order to improve the practice of addressing misconceptions. Finally, patterns of change illustrated significant improvements in all five components of reformed instruction only after all 3 years of treatment, when examining the treatment group over time. This again reveals the need for 3 years of sustained professional development for improving the practice of reformed science education.

These findings show patterns of change over time and further support the hypothesis that sustained professional development affects science instruction over time. Two years of professional development is somewhat effective in changing some components of reformed instruction; however, sustained professional development of 3 years is necessary in order to significantly improve all five components of reformed science education concomitantly.

Interactions between overall science instruction and the specific components. Finally, statistical analysis of the data revealed how instruction changes over time in regard to interaction between performance in overall classroom science instruction and the practice of specific components of reformed science instruction of elementary school teachers. Positive linear relationships were found between the Summary Judgment of Instruction scores, the PESTL Total scores, and the specific components of reformed science instruction scores of elementary school teachers over time. The gradual positive trend illustrated for the Summary Judgment of Instruction scores was similar to that of content alignment and modeling scores, whereas the trend of investigation illustrated greater positive direction. Addressing misconceptions and talk and argument scores illustrated a step-like line segment similar gains over time with the PESTL total score. All three of these measures showed a flat trend in year one and a significantly increased positive trend during year two and three of treatment (see Figure 4.1).

Mention should be made regarding the high correlation between improvements in Talk and argument and the overall PESTL Observation Protocol rating. This may be explained by the higher number of indicators in the talk and argument component of the PESTL Observation Protocol. As mentioned previously, there were five indicators with a possible score of 25 for Talk and argument compared to three or less indicators in the other four components. Therefore, talk and argument scores are more likely to influence the PESTL Total scores.

These findings show interactions between improvement in overall science instruction and the components of reformed science instruction over the 3 years of professional development. These findings further support the hypothesis that teachers who receive sustained professional development will improve in overall science instruction and specific components of reform science education over time, and that interactions between measures will occur.

Transformational learning theory and question three findings. Transformational learning theory helps explain and inform the findings from question three in examining comparisons, patterns, and interactions in overall and specific components of reformed science instruction. In relation to the transformational learning theory, findings for question three suggest that PESTL professional development balanced the two modes of transformational learning (Mezirow & Taylor, 2009) when comparing the components of reformed science instruction. The most influenced components of investigation (instrumental), talk and argument (communicative), modeling (instrumental) and addressing misconceptions (communicative) show an even distribution of instrumental and communicative modes.

Patterns of change over time also illustrated the concepts of the transformational learning theory. As mentioned earlier, transformational learning theory suggests that adult learning takes time (Mezirow, 2000). This concept is especially applicable when examining patterns of change over time. Findings from this study suggest that 3 years of professional development were necessary to transform overall and all five components of reformed science educations concomitantly.

Interactions between changes in overall instruction and changes in practices of the specific components of reformed science education further illustrated the concepts of the transformational learning theory. Findings suggest that a balance between instrumental and communicative learning sustained over a 3-year period have synergistic effects. In relation to the transformational learning theory it appears that the balance of instrumental and communicative modes in sustained professional development activities in the PESTL program led to changes in classroom instruction that were both instrumental and communicative in nature. Findings suggest that positive interactions between overall instruction scores and the specific components of reformed science education may occur as a result of balancing instrumental and communicative learning in sustained professional development.

Implications of Study

The following conclusions can be drawn about the impacts of sustained professional development on the classroom science instruction of elementary school teachers.

- Teachers who participate in sustained professional development, which is designed to meet recommendations for reform in elementary science education, are likely to have significantly more effective overall science instructional practices.
- Teachers who participate in sustained professional development, which is designed to meet recommendations for reform in elementary science

education, are likely to have significantly higher scores in specific components of reformed science instruction.

- Transformation of the overall science instruction of elementary school teachers may take place when teachers participate in sustained professional development for at least 2 years.
- Concomitant transformation of all five specific components of reformed science instruction of elementary school teachers is more likely to take place when teachers participate in sustained professional development for at least 3 year.
- One year of professional development does not appear to significantly improve the practices of overall science instruction or specific components of reformed science instruction of elementary teachers who participate when compared to nonparticipants.
- Two years of professional development were necessary in order to significantly improve the practices of science talk and argument, investigation, and content alignment in the classroom instruction of elementary teachers who participate when examining treatment group scores over time.
- Three years of professional development were necessary in order to significantly improve the practices of modeling and addressing misconceptions in the classroom instruction of elementary teachers when examining treatment group scores over time.

- Professional development, sustained for 3 years, is effective in simultaneously transforming the overall science instruction and all five components of reformed science instruction of participating elementary school teachers.

The PESTL professional development program impacted teachers' overall science instruction as well as components of reformed science instruction. Teachers' scores in all seven outcome measures rose significantly after 3 years of professional development.

Limitations of Study

There were a number of limitations to this study. They include the following.

- Though the treatment population size is large, the sample size is small, so generalization may not be made. In future studies this limitation could be addressed by following a large number of participants or the entire population of participants.
- Because the treatment group was not observed prior to treatment, the control group must serve to illustrate a baseline for treatment. This is not ideal, but measures were taken to ensure the data set met requirements to serve this purpose. This is a threat to the validity of findings, because the control group may not have performed the same as the treatment group prior to treatment. In future studies this limitation could be addressed by observing participants prior to treatment.
- The control group was not followed over the 3 year, so a comprehensive repeated measures model could not be used between the treatment and control

groups across time, instead independent *t* tests were used to compare the control group to the treatment group separately. This is a threat to the validity of findings because no direct comparison can be made between the control group and treatment group over time. In future studies this limitation could be addressed by observing the control teachers prior to treatment and over time, then direct comparisons could be made.

- Though measures were taken to limit the contamination of the control group, it is possible that some diffusion effects occurred with treatment teachers associated with control teachers outside of the school setting. This is a threat to the validity of findings because control teachers may have been exposed to treatment ideas in associations outside the school setting. This limitation is difficult to address in educational studies because isolation of control teachers is not realistic. In future studies this limitation could be reduced by only measuring instruction specific to treatment. For example, the content alignment component is not specific to reformed science education; therefore, this component could be changed to a practice more targeted by professional development designed to reform science instruction.
- The control group was convenience sampled, not stratified random sample. With the data set being preexisting this limitation could not be avoided. This is a threat to the validity of findings because convenience samples may not represent the population of nonparticipants. In future studies this limitation could be addressed if control groups were stratified random sampled from all

teachers in the districts who are not participating.

- Some teachers were more experienced than others and some had more science background than others. Because all treatment and control teachers had a general elementary teaching certification, scientific background and exposure were not considered. Other than grade level, teacher information was not gathered on participants or control teachers. This is a threat to the validity of findings because teaching experience and science background may have influenced instruction. In future studies this limitation could be addressed by gathering information about years of teaching experience and science background on the treatment and control teachers.
- Participants varied in age and personal responsibilities. Some may have had more time constraints due to school or family obligations, while others may have had more free time to engage in greater preparation for classroom science instruction. As mentioned previously, teacher information was not gathered on participants or control teachers. This is a threat to the validity of findings because teacher age and personal responsibilities may have influenced instruction. In future studies this limitation could be addressed by gathering information about age and personal responsibilities on the treatment and control teachers.
- Entire schools were selected for PESTL treatment but individual teacher participation was voluntary. This situation, however, is typical of inservice professional development programs and is difficult to avoid in professional

development research. Refusal rate of teachers at each school was not recorded. In future studies this limitation could be addressed by gathering information about the teachers at the school and recording refusal rates and possibly the reason for refusal of each teacher not participating at the school.

- Finally, this researcher as well as the PESTL program director were two of the six science experts involved in data collection. The use of quantitative data and the triangulation of data sources served to decrease the observer bias that may have developed over the course of the study; however, it should be recognized that the presence of the director as an observer may have influenced the teachers in terms of implementing more reformed practices than if the research and the classroom observations had not been conducted. The Cronbach's alpha reliability for the Summary Judgment of Effective Instruction tool used for this study was established at .83, and the Cronbach's alpha reliability for the PESTL Observation Protocol tool was established at .84. These rater reliability scores are high and support the methods for rating instruction are beyond the subjective judgments of the observers.

Recommendations for Practice

Science professional developers, teacher educators, school principals, district administrators, and science education policy makers may apply the results of this study to improve practice in elementary science teacher development. This study supports the findings of multiple studies on improving elementary science instruction. These studies suggest the need for professional development to be long-term, embedded in teaching practice in the classroom, and rooted in research on how children learn science (Duschl et al., 2006; van Driel et al., 2001; Weiss et al., 2003; Yager, 2000). Educational leaders, who desire to improve classroom science instruction, may use these findings to support their decision to provide sustained science professional development.

Nationally, a concentrated effort has been made to reform science education through teacher professional development; however, most of the programs are short term and fail to offer teachers the science literacy and continued support they need in order to transform the science teaching and learning in their classrooms (Birman et al., 2000; Duschl et al., 2006; Elmore, 2002). This study suggests that sustained professional development, which is designed to meet recommendations for reform in elementary science education, is likely to improve the science classroom instruction of participating elementary school teachers over time. Based on the findings of this study, teacher development programs in elementary science education should incorporate the following.

- Sustain elementary science professional development for at least 2 years, and ideally 3 year, in order to transform the classroom science instruction of elementary school teachers.
- Design teacher development programs based on recommendations for reform in elementary science education.
- Design teacher development to be embedded in teaching practice in the classroom and rooted in research on how children learn science.
- Honor teachers as adult professionals who are capable of transforming their

understanding of science teaching and learning.

- Show respect for the participating teachers' background knowledge and experiences in science, addressing misconceptions just as they are asked to do for the students in their classrooms.
- Design teacher development programs to include partnerships between teachers, schools, science teacher educations programs, professional developers, science experts, science organizations, and university science departments to provide the science literacy and continued support teachers need in order to reform the science teaching and learning in their classrooms.

Professional developers across the U.S. are implementing such programs (Moulding, 2008; NCES, 2009, 2011). According to this study, sustained professional development programs designed to meet recommendations for reform in elementary science education are likely to be effective in transforming science instruction in the elementary classroom.

Researcher Reflections

The researcher learned much conducting this study. Some attention should be given to the researcher's bias, assumptions, experiences, education, and preconceived ideas. The researcher is a past elementary teacher and current college elementary teacher educator. The researcher's science background includes 2 years of graduate-level study and research in the life sciences, nine years of teaching integrated science to early childhood and elementary school children, 2 years of teaching K-8th-grade children at a local nature center, and 2 years teaching the methods of science in a college elementary teacher education program.

Throughout the study the researcher tried to remain as an outside observer, separated from the PESTL program participants, facilitators, and director. This allowed her to remain neutral in regard to outcomes and data collection. However, after studying the research, reading the literature on recommendations for professional development and spending 2 years studying and observing the PESTL professional development program, the researcher did make assumptions that the program would be effective in changing instruction. Though the researcher commenced the study absent of preconceived ideas about the program, study and experience in the program lead to biases in favor of the program's effectiveness even before the statistical analyses were completed.

This researcher was changed by conducting the study and has found a new passion for sustained teacher development in elementary science education. This study found many strengths, yet also exposed some weaknesses in the PESTL program. Through detailed examination of the PESTL observation tools and data collection procedures, areas of improvement were identified. To help improve the PESTL program the researcher has been invited to contribute to the new round of PESTL professional development, which started in the summer of 2011.

Suggestions for Future Research

Several future avenues for research are suggested by this study. These include the following.

- *Study continuation*. How will changes in study participants' science instruction evolve in future years after the sustained professional development is over? A quantitative study, using the same observation instruments, could follow participants for up to 3 years after treatment to determine the effects sustained professional development has on future practice.

- *Teacher science literacy*. What is the impact of sustained professional development on teacher science literacy? A quantitative study could examine the changes in science literacy over time by administering a science literacy test before treatment, during, and after treatment. A control group could also be followed to make comparisons.
- *Teacher efficacy*. What is the impact of sustained professional development on teacher-efficacy in science instruction? A quantitative or qualitative study could examine the changes in teacher-efficacy over time by administering self-efficacy tests, questionnaires, journals, and/or interviews before treatment, during, and after treatment. A control group could be followed to make comparisons. This could also be done as mixed-methods study.
- *Teachers' demographic beliefs*. What impact can sustained science professional development have on teachers' beliefs in the abilities of their students to learn science? A qualitative study could examine the changes in teachers' beliefs about the ability of their students to learn scientific concepts and their understanding of how children learn science over time. Questionnaires, journals, and/or interviews could be used to gather information for this type of study.

- *Impact of professional development on student achievement*. What is the impact of sustained professional development on students' science standardized test (CRT) scores? A quantitative study could examine the changes in student science achievement over time by analyzing science CRT test scores before treatment and over time during treatment. A control group could also be followed to make comparisons.

Future research into these topics could contribute to knowledge of the effects of sustained professional development on elementary science education and may help future professional developers design and implement even more effective professional development programs. Such programs would lead to successful reform in science education, including transformed teacher beliefs and practices, and increased student achievement and science understanding.

Conclusion

This quantitative quasi-experimental study explored the effects of sustained professional development, which was designed to meet recommendations for reform in science education, on the classroom science instruction of elementary school teachers. The theoretical framework proposed that elementary science professional development is linked to transformational learning (Mezirow, 2000) as inservice teachers transform their taken-for-granted frames of reference about science teaching in order to make them more inclusive, discriminating, reflective, and consistent with recommendations for reformed science education.

The literature established connections between professional development in elementary science education and the desired results of reformed classroom instruction. It established transformative learning theory as the theoretical lens for studying professional development in elementary science education. The review of literature identified professional development characteristics proven significant in improving classroom science instruction of elementary school teachers. Additionally, five components of effective instruction in elementary science education were identified and described. The five components included: (a) talk and argumention, (b) modeling and representations, (c) investigations and inquiry, (d) alignment to science core concepts, and (e) appropriately addressing science misconceptions. Recent studies on the effects of teacher development on classroom instruction were explored, analyzed, and compared to this study. Finally, an overview of the PESTL professional development was given. This sustained professional development in elementary science education was the program under examination in this study.

The effectiveness of science instruction was measured through classroom observations. There were 22 treatment participants and 20 control teachers in this study. Treatment teachers participated in all 3 years of PESTL professional development and were observed each year. Control teachers did not participate and were observed once during the 3 years of the study. The control group scores served as a baseline for nontreatment. In order to determine the changes in classroom instruction two observation data sets were used. One set was data gathered from the Summary of Effective Science Instruction tool. The other set was data gathered from the PESTL Observation Protocol tool. With the PESTL Observation Protocol data, five components of reformed science instruction were analyzed. The five components were: (a) talk and argument, (b) investigation, (c) modeling, (d) content alignment, and (e) addressing misconceptions. The data analyses utilized were independent *t* test and one-way factorial ANOVAs. Independent *t* tests with Bonferroni correction were used for examining data between the control group and treatment group. One-way repeated measures factorial ANOVAs with Bonferroni post-hoc were used for examining change over time within the treatment group.

The findings illustrated that teachers, who participated in PESTL professional development, had significantly higher scores in Overall Summary of Science Instruction and PESTL Observation Protocol scores. Findings further illustrated that there were most and least influenced components of PESTL, patterns of change among PESTL components, and that interactions between overall instruction and PESTL components did exist. These findings support the hypothesis that third- through sixth-grade teachers who receive sustained professional development will likely have improved effectiveness in classroom science instruction.

Implications of this study are that transformation of overall science instruction and the five components of reformed science instruction of elementary school teachers is likely to take place when teachers participate in well-designed sustained professional development programs. 3 years of sustained professional development is ideal for significantly improving overall science instruction and all five components of reformed science instruction concomitantly. In sum, the sustained professional development program impacted teachers' overall science instruction as well as the five components of reformed science instruction. Teachers' scores in all six outcome measures rose significantly over the 3 years of professional development.

REFERENCES

Abd-El-Khalich, F., & Akerson V. L. (2004). Learning as conceptual change: Factors mediated the development of preservice elementary teachers' views of nature of science. *Science Education, 88,* 785-810.

Akerson, V. L. (2005). How do elementary teachers compensate for incomplete science content knowledge? *Research in Science Education, 35,* 245-268.

American Association for the Advancement of Science. (1989). *Science for all Americans: A Project 2061 report on literacy goals in science, mathematics, and technology*. Washington, DC: Author.

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy.* New York, NY: Oxford University Press.

American Educational Research Association. (2005). *Studying teacher education: The report of the AERA panel on research and teacher education* (M. Cochran-Smith & K. M. Zeichner, Eds.). Mahwah, NJ: Erlbaum.

American Federation of Teachers. (2002). *Principles for professional development* (Report 39-0176). Washington, DC: Author.

Appleton, K. (2003). How do beginning primary school teachers cope with science? Toward an understanding of science teaching practice. *Research in Science Education, 33,* 1-25.

Association for Supervision and Curriculum Development. (2007). *On site staff development*. Retrieved from http://webserver3.ascd.org/ossd/planning.html

Barnes, D. (1992). *From communication to curriculum* (2nd ed.). Portsmouth, NH: Boynton/Cook.

Berk, L., & Winsler, A. (1995). *Scaffolding children's learning: Vygotsky and early childhood education.* Washington, DC: National Association for the Education of Young Children.

Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education, 94,* 765-793.

Birman, B., Desimore, L., Porter, A., & Garet, M. (2000). Designing professional development that works. *Educational Leadership, 57*(8)*,* 28-33.

Briggs, J. (1987). *Student approaches to learning and studying.* Hawthorn, Australia: Australian Council for Educational Research Ltd., Radford House.

Bruning, R., Schraw, G., & Ronning, R. (1999). *Cognitive Psychology and instruction* (3rd ed.). Upper Saddle River NJ: Prentice Hall.

Buckley, B. C., Gobert, J. D., Kindfield, A. C. H., Horwitz, P., Tinker, R. F., Gerlits, B., … Wilensky, U. (2004). Model-based teaching and learning with BioLogica: What do they learn? How do they learn? How do we know? *Journal of Science Education and Technology, 13,* 23-41.

Campbell, T. (2012). Building community in triads involved in science teacher education: An innovative professional development model. *Brock Education Journal, 21*(2)*,* 1-17.

Carlone, H., & Webb, S. (2006). On (not) overcoming our history of hierarchy: Complexities of university/school collaboration. *Science Education, 90,* 544-568.

Carey, S. (2000). Science Education as conceptual change. *Journal of Applied Developmental Psychology, 21,* 13-19.

Chaiklin, S. (2001). *The theory and practice of cultural-historical psychology.* Aarhus, Denmark: Aarhus University Press.

Chi, M. T. H. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In S.Vosniadou (Ed.), *Handbook of research on conceptual change* (pp. 61-82). Hillsdale, NJ: Erlbaum.

Cohen, L., & Manion, L. (2000). *Research methods in education* (5th ed.). London, England: Routledge.

Corcoran, T. B., McVay, S., & Riordan, K. (2003). *Getting it right: The MISE approach to professional development* (Research Report No. RR-055). Philadelphia, PA: Consortium for Policy Research in Education, University of Pennsylvania.

Craven, R., Marsh, H., Debus, R., & Jayasinghe, U. (2001). Diffusion effects: Control group contamination threats to the validity of teacher-administered interventions. *Journal of Educational Psychology, 93,* 639-645.

Crawford, B. A. (2000). Embracing the essence of inquiry: New roles for science teachers. *Journal of Research in Science Teaching, 37,* 916-937.

Cuevas, J., Lee, O., Hart, J. E., & Deaktor, R. (2005). Improving science inquiry with elementary students of diverse backgrounds. *Journal of Research in Science Teaching, 42,* 337-357.

Darling-Hammond, L., & Baratz-Snowden, J. (2007). A good teacher in every classroom: Preparing the highly qualified teachers our children deserve. *Educational Horizons, 85,* 111-132.

Darling-Hammond, L., & McLaughlin, M. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappan, 77,* 597-603.

Davis, K. S. (2003). Change is hard: What science teachers are telling us about reform and teacher learning of innovative practices. *Science Education, 87,* 3-30.

Dawes, L. (2004). Talk and learning in classroom science. *International Journal of Science Education, 26,* 677-695.

diSessa, A., & Minstrell, J. (1998). Cultivating conceptual change with benchmark lessons. In J. Greeno & S. Goldman (Eds.), *Thinking practices* (pp. 155-187). Hillsdale, NJ: Erlbaum.

Driel, F. H. V., & Verloop, N. (1999). Teachers' knowledge of models and modeling in science. [Electronic version]. *International Journal of Science Education, 21,* 1144-1153.

Drits, D., & Stark, L. (2011, April).g *The role of school context in advancing teacher change in the year following a science-based professional development program.* Paper presented at the AERA Annual Meeting, New Orleans, LA.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education, 84,* 287-312.

DuFour, R., Eaker, R., & Many, T. (2006). *Learning by doing: A handbook for professional learning communities at work*. Bloomington, IN: Solution Tree.

Duschl, R., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education, 38*, 39-72.

Duschl, R., Schweingruber, H. A., & Shouse, A. W. (2006). *Taking science to school: Learning and teaching science in grades K-8.* Washington, DC: The National Academy Press.

Elmore, R. F. (2002). *Bridging the gap between standards and achievement: The imperative for professional development in education*. Washington, DC: Albert Shanker Institute.

Enoch, L., & Riggs, I. (1990). Further development of an elementary science teaching efficacy belief instrument: A preservice elementary scale. *School Science and Mathematics, 63,* 694-706.

Fitch, T., & Fisher, R. (1979). Survey of science education in a sample of Illinois schools: Grades K-6. *Science Education,* 63 (3) 407-416

Forgione, P. (2006). *International test scores: Poor U.S. test results tied to weak curriculum*. Retrieved from http://4brevard.com/choice/international-test-scores.htm

Franco, C., & Colinvaux, D. (2000). Grasping mental models. In J. K. Gilbert & C. J. Boulter (Eds.), *Developing models in science education* (pp.93-118). Dordrecht, The Netherlands: Kluwer.

Garet, M., Birman, B. F., Porter, A. C., Desimone, L., Herman, R., & Yoon, K. S. (1999). *Designing effective professional development: Lessons from the Eisenhower program.* Washington, DC: American Institutes for Research.

Garet, M., Porter, A., 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. (1989). What is literacy? *Journal of Education, 171*(1), 18-25.

Giere, R. (1991). *Understanding scientific reasoning.* Orlando, FL: Holt, Rinehart, and Winston.

Gilbert, J. (1993, April). *The role of models and modeling in science education*. Paper presented at the 1993 Annual Conference of the National Association for Research in Science Teaching, Atlanta, GA.

Gilbert, J. (1995, April). The role of models and modeling in some narratives in science learning. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Gilbert, J., & Boulter, C. (1998). Learning science through models and modelling. In B. Fraser & K. Tobin (Eds), *International handbook of science education* (pp. 52-66) Dordrecht, The Netherlands: Kluwer.

Gilbert, J. K., Boulter, C. J., & Elmer, R. (2000). Positioning models in science education and in design and Technology education. In J. K. Gilbert & C. J. Boulter (Eds.), *Developing models in science education* (pp.3-17). Dordrecht, The Netherlands: Kluwer.

Gilbert, S. (1991). Model building and a definition of science. *Journal of Research in Science Teaching, 28,* 73-79.

Glickman, C. D., Gordon, S. P., & Ross-Gordon, J. M. (2009). *The basic guide to supervision and instructional leadership* (2nd ed.). Boston, MA: Pearson.

Gobert, J. D., & Buckley, B. C. (2000). Introduction to model-based teaching and learning in science education. *International Journal of Science Education, 22,* 891-894.

Greca, I. M., & Moreire, M. A. (2000). Mental models, conceptual models, and modeling. [Electronic version]. *International Journal of Science Education*, 1, 1-11. Retrieved from www.tandfonline.com/doi/pdf/10.1080/095006900289976

Hassard, J., & Dias, M. (2008). *The art of teaching science: Inquiry and innovation in middle school and high school* (2nd ed.). London, England: Routledge.

Hazen, R. M., & Trefil, J. (2009). *Science matters: Achieving scientific literacy* (2nd ed.) New York, NY: Anchor Books.

Hiebert, J. (1997). *The third international mathematics and science study (TIMSS): A comprehensive analysis of elementary and secondary math and science education*. Retrieved from http://www.house.gov/science/hiebert_10-8.htm

Hodson, D. (2008). *Towards scientific literacy: A teachers guide to the history, philosophy and sociology of science.* Rotterdam, The Netherlands: Sense Publishers.

Horizon Research, Inc. (2001). *Evaluation manual: Classroom observation protocol*. Chapel Hill, NC: Author.

Hurd, P. D. (1958) Science literacy: Its meaning for American schools. *Educational Leadership*, 16(1), 13-16.

Ingham, A. M., & Gilbert, J. K. (1991). The use of analogue models by students of chemistry at higher education level. *International Journal of Science Education, 13,* 193-202.

International Association for the Evaluation of Educational Achievement (IEA), Trends in International Mathematics and Science Study (TIMSS). (1995). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

International Association for the Evaluation of Educational Achievement (IEA),Trends in International Mathematics and Science Study (TIMSS). (2003). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

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

Kennedy, M. (1998). Education reform and subject matter knowledge*. Journal of Research and Science Teaching, 35,* 249-263.

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

Knapp, M. S., McCaffrey, T., & Swanson, J. (2003, April). *District support for professional learning: What research says and has yet to establish.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Kukla, A. (2000). *Social constructivism and the philosophy of science*. London, England: Routledge.

Labov, J. B. (2006). National and state standards in science and their potential influence on undergraduate science education. *Life Sciences Education, 5,* 204-209.

Lehrer, R., & Schauble, L. (2004). Modeling natural variations through distribution. *American Educational Research Journal, 41,* 635-679.

Lehrer, R., & Schauble, L. (2006). Cultivating model-based reasoning in science education. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 371-387). Cambridge, MA: Cambridge University Press.

Lemke, J. L. (1990). *Talking science: Language, learning, and values*. Norwood, NJ: Ablex.

Lemke, J. L. (2001). Articulating communities: Sociocultural perspectives on science education. *Journal of Research in Science Teaching, 38,* 296-316

Linn, M. C., & Muilenberg, L. (1996). Creating lifelong science learners: What models form a firm foundation. *Educational Researcher, 25*(5), 18-24.

Lohr, S. L. (1999). *Sampling: Design and analysis*. Belmont, England: Duxbury.

Loucks-Horsley, S. (1998). The role of teaching and learning in systemic reform: A focus on professional development. *Science Educator, 7,* 1-6.

Loucks-Horsley S., Love, N., Stiles, K. E., Mundry, S., & Hewson, P. W. (2003). *Designing professional development for teachers of science and mathematics* (2nd ed.). Thousand Oaks, CA: Corwin.

McCurdy, R. C. (1958). Towards a population literate in science. *The Science Teacher*, *25,* 366-368.

McIntyre, D., & Hagger, H. (1992). Professional development through the Oxford Internship Model. *British Journal of Educational Studies, 40,* 264-283.

Metz, K. E. (1995). Reassessment of developmental constraints on children's science instruction. *Review of Educational Research, 65,* 93-127.

Mezirow, J. (1991). *Transformative dimensions of adult learning*. San Francisco, CA: Jossey-Bass.

Mezirow, J. (2000). *Learning as transformation: critical perspectives on a theory in progress*. San Francisco, CA: Jossey Bass.

Mezirow, J., & Taylor, E. W. (2009). *Transformative learning in practice: Insights from community, workplace, and higher education.* San Francisco, CA: Jossey Bass.

Michaels, S., Shouse, A. W., & Schweingruber, H. A. (2008). *Ready, set, science! Putting research to work in K-8 science classrooms.* Washington, DC: NA Press.

Minuskin, S. O. (2009). *Effects of professional development on the knowledge and classroom practices of elementary school science teachers* (Doctoral dissertation). Available from ProQuest Dissertation and Theses database. (UMI No. 3385691)

Moulding, B. D., & Baird, H. (2008, January). *Utah science core curriculum and ready, set, SCIENCE!* Paper presented to State Science Education Coordination Committee. Salt Lake City, UT.

National Assessment of Educational Progress. (2009). Retrieved from http://nces.ed.gov/ pubsearch/pubsinfo.asp?pubid=2011451

National Center for Education Statistics. (2001). *Teacher quality: A report on the preparation and qualifications of public school teachers.* Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

National Center for Education Statistics. (2009). *Teacher quality: A report on the preparation and qualifications of public school teachers*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.

National Research Council. (1999). *How people learn: Brain, experience and school*. Washington, DC: National Academy Press.

National Research Council. (2005). *How students learn: Science in the classroom*. M. S. Donovan & J. R. Bransford (Eds.). Washington, DC: National Academy Press.

National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8.* Washington, DC: Committee on Science Learning, Kindergarten Through Eighth Grade.

National Research Council. (2011). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas.* Washington, DC: The National Academies Press.

National Science Teachers Association. (2003). *Standards for science teacher preparation NSTA*. Retrieved from http://www.nsta.org/main/pdfs/ NSTAstandards2003.pdf

National Staff Development Council. (2001). *NSDC standards for staff development*. Oxford, OH: Author.

Neuman, S. (2009). *Changing the odds for children at risk: Seven essential principles of educational programs that break the cycle of poverty.* Westport, CT: Praeger.

Norman, D. A. (1983). Some observations on mental models. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 7-14). Hillsdale, NJ: Erlbaum.

Odden, A. R., Archibald, S., Fermanich, M. L., & Gallagher, H. A. (2002). A cost framework for professional development. *Journal of Education Finance*, *28*(1), 51-74

Ornek, F. (2008). Models in science education: Applications of models in learning and teaching science. *International Journal of Environmental & Science Education, 3*(2), 35-45.

Osborne, M. D. (1998). Teacher as knower and learner: Reflections on situated knowledge in science teaching. *Journal of Research in Science Teaching, 35,* 427-439.

Parsad, B., Lewis, L., & Farris, E. (2001). *Teacher preparation and professional development: 2000*. Washington, DC: U.S. Department of Education.

Peacock, A., & Rawson, B. (2001). Helping teacher develop competence criteria for evaluating their professional development. *International Journal of Educational Development, 21,* 79-92.

Pella, M. O., O'Hearn, G. T., & Gale, C. W. (1966). Referents to scientific literacy. *Journal of Research in Science Teaching, 4,* 199-208.

Perkins, D. N. (1986). *Knowledge as design.* Hillsdale, NJ: Erlbaum.

Petrosino, A., Lehrer, R., & Schauble, L. (2003). Structuring error and experimental variation as distribution in the fourth grade. *Mathematical Thinking and Learning, 5,* 131-156.

Piaget, J., & Inhelder, B. (1969). *The osychology of the child.* New York, NY: Basic.

Piburn, M. D., Sawada, D., Falconer, K., Turley, J., Benford, R., & Bloom, I. (2000). *Reform teaching observational protocol handbook*. Tempe, AZ: Arizona Board of Regents.

Richardson, M. R., & Liang, L. L. (2008). The use of inquiry in the development of preservice teacher efficacy in mathematics and science. *Journal of Elementary Science Education, 20,* 1-16.

Riggs, I. M. (1991, April). *Gender differences in elementary science teacher self-efficacy.* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

Rutherford, F. J., & Ahlgren, A. (1990). *Science for all Americans.* New York, NY: Oxford University Press.

Santau, A. O. (2008). *Elementary teachers' knowledge and practices in teaching science to English language learners* (Doctoral dissertation). Retrieved from http://scholarlyrepository.miami.edu/oa_dissertations/294

Schneider, R., & Krajcik, J. (2002). Supporting science teacher learning: The role of educative curriculum materials. *Journal of Science Teacher Education, 13,* 221-245.

Sottile, J. M., Carter, W., & Murphy, R. A. (2002, April). *The Influence of self-efficacy on school culture, science achievement and math achievement among inservice teachers*. Paper presented at the Annual meeting of the American Education Research Association, New Orleans, LA.

Sparks, D., & Hirsh, S. (1997). *A new vision for staff development*. Alexandria, VA: Association for Supervision and Curriculum Development.

Stein, M., Smith, M., & Silver, E. (1999). The development of professional developers: Learning to assist teachers in new settings in new ways. *Harvard Educational Review, 69,* 237-269.

Stigler, J., & Hiebert, J. (1999). *The teaching gap*. New York, NY: The Free Press.

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,* 963-980.

Taylor, E. W. (2000). Analyzing research on transformative learning theory. In J. Mezirow (Ed.), *Learning as transformation: Critical perspectives on a theory in practice* (pp. 285-328). San Francisco, CA: Jossey-Bass.

Tharp, R. G., & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning, and schooling in social context.* Cambridge, England: Cambridge University Press.

Tippett, C. (2009). Argumentation: The language of science. *Journal of Elementary Science Education, 21*(1), 17-25.

Thorson, A. (2002). *By your own design: A teacher's professional learning guide.* Retrieved from http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED461517&ERICExtSearch_SearchType_0=no&accno=ED461517

Toulmin, S. E. (1958). *The uses of argument*. London, England: Cambridge University Press.

Tregidgo, D., & Ratcliffe, M. (2000). The use of modeling for improving pupils' learning about cells. *School Science Review, 81,* 53-59.

Utah State Office of Education. (2002). *Utah elementary science core curriculum.* Retrieved from www.usoe.org

van Driel, J. H., Beijaard, D., & Verloop, N. (2001). Professional development and reform in science education: The role of teacher's practical knowledge. *Journal of Research in Science Teaching*, *38*, 137-158.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning & Instruction, 4,* 45-69.

Vosniadou, S. (2004). Extending the conceptual change approach to mathematics learning and teaching. *Learning & Instruction, 14*, 445-451.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Vygotsky, L. (1986). *Thought and language.* Cambridge, MA: The MIT Press.

Weiss, I. R. (1994). *A profile of science and mathematics education in the United States: 1993.* (Report No. RTI/2938/00-FR). Chapel Hill, NC: Horizon Research. (Eric Document Reproduction service No. ED 382 461)

Weiss, I. R., Pasley, J. D., Smith, P. S., Banilower, E. R., & Heck, D. J. (2003). *Looking inside the classroom: A study of K-12 mathematics and science education in the United States.* Chapel Hill, NC: Horizon Research.

Wenger, E. (1998). *Communities of practice: Learning, meaning and identity*. Cambridge, MA: Cambridge University Press.

Wertsch, J. V. (1998). *Mind as action*. New York, NY: Oxford University Press.

Wu, H., & Krajcik, J. S. (2006). Inscriptional practices in two inquiry-based classrooms: A case study of seventh graders' use of data tables and graphs. *Journal of Research in Science Teaching, 43,* 63-95.

Yager, R. (2000). A vision for what science education should be like for the first 25 years of a new millennium. *School Science and Mathematics, 100,* 327-41

Yore, L., Bisanz, G., & Hand, B. (2003). Examining the literacy component of science literacy: 25 years of language arts and science research. *International Journal of Science Education, 25*, 689-725.

Zemelman, S., Daniels, H., & Hyde, A. (2005). *Best practice: Today's standards for teaching and learning in American schools* (3rd ed.). Portsmouth, NH: Heinemann.

APPENDICES

Appendix A

Copyright Release Permission and Summary Judgment of Science Instruction

July 15, 2012

Nancy Hauck
886 E. 970 S. Circle
St. George, UT 84790
hauck@dixie.edu

Dear Brett Moulding,

I have prepared my dissertation in the Teacher Education and Leadership Department spring semester 2012 at Utah State University. I completed my degree in May of 2012.

The dissertation, The Effects of Sustained Effects of Sustained Teacher Professional Development on the Classroom Science Instruction of Elementary School Teachers, of which I am the sole author includes the Summary Judgment of Science Instruction and PESTL Observation Protocol of Instructional Effectiveness in the appendices, of which is copyrighted to Essential Teaching and Learning, LLC.

As you are the Permission Editor of Essential Teaching and Learning, LLC and PESTL Director, I would like permission to include the Summary Judgment of Science Instruction and PESTL Observation Protocol of Instructional Effectiveness in my appendices. Please note that USU sends dissertation to Bell & Howell Dissertation Services to be made available for reproduction.
Copyright and permission information will be included in in the appendices. If you would like a different acknowledgment, please so indicate.

Please indicate your approval of this request by signing in the space provided, and attach any other form necessary to confirm permission. If you have any questions, please contact me at the email above.

Thank you for your assistance.
Nancy Hauck

I, Brett Moulding, hereby give permission to Nancy Hauck to include the requested documents in her dissertation with the following acknowledgment:

PESTL Science Classroom Observation Protocol – Copyrights are exclusive property of Essential Teaching and Learning, LLC.

Signed Brett Moulding

Date July, 16, 2012

Summary Judgment of Science Instruction

Use the rubric below to identify the level of instruction.

Circle one

I-a or I-b	II	III	IV	V

Level 1 There is little evidence of student thinking or engagement with important ideas of mathematics/science. Instruction is **not likely** to enhance students' understanding of big ideas or concepts.

a) Passive "Learning" – Instruction is pedantic and uninspiring. Students are passive recipients of information from the teacher or textbook; material is presented without scaffolding for students.
b) Activity for Activity's Sake – Students are involved in hands-on activities or other individual or group work, but it appears to be activity for activity's sake. Lesson lacks a clear sense of purpose and/or a clear link to conceptual development.

Level 2 Instruction includes some elements of effective practice, but improvement is needed (e.g. content not aligned to core, student learning difficulties are ignored, teacher does not check for understanding). The lessons will **not likely** lead students to understanding important science concepts.

Level 3 Instruction is purposeful and characterized by quite a few elements of effective practice. Content is well aligned to the Core, but the ILOs are not featured within the lesson or ILOs are the focus and science content is missing. Students are engaged in meaningful activities, but instruction **does not focus** on big ideas or use student engagement in thinking and making connections.

Level 4 Instruction is purposeful and engaging for most students. Students actively participate in meaningful work (e.g., investigations, teacher/instructor presentations, discussions with each other or the teacher/instructor, reading). The lesson is well-designed and aligned to the Core. The instruction will **likely** lead to meaningful student learning.

Level 5 Instruction is purposeful and all students are highly engaged most of the time in meaningful science learning (e.g., investigation, teacher/instructor presentations, discussions with each other or the teacher/instructor, reading). The lesson is well-designed, aligned to the Core and the teacher's craft is implemented, with flexibility and responsiveness to students' needs and interests. Instruction is **very likely to lead** to students' understanding of science/math concepts, skills and processes. The content and ILOs of the Core are learned and applied.

*PESTL Science Classroom Observation Protocol – Copyrights are exclusive property of Essential Teaching and Learning, LLC.

Appendix B

PESTL Observation Protocol

PESTL Observation Protocol

Section II: INSTRUCTION

1) Instructional Activity

a) Subject/Title	
b) *Source	
c) Description of the Instructional Activity:	

2) Observation of Instructional Activity (45 minutes *minimum of 30 minutes)

Instructional time period observed (# of minutes)				
* What portion of the planned instructional activity was observed?		%		
Instructional Effectiveness	**0-5**	**Comments/Tally Marks**		
Talk and Argument Classroom Discussion				
a) Student/Teacher interaction ratio		Q-1	Q-2	Q-3
b) Number of times teacher actively extends thinking				
c) Number of examples and analogies in presentations				
d) Supports relevant inter-student discussion				
e) Uses accurate science language				
Sub Total				
Investigation				
f) Science investigations are directed by teacher		formulate question, make observation, formulate a testable question, identify variable, record data, make inferences, classify, discuss limitations of findings, analyze data, inquire		
g) Science investigations are student centered and in small groups				
h) Science concepts within investigation are assessed				
Sub Total				

Modeling		
i) Uses models to demonstrate concepts		
j) Uses models to assess student understanding		
k) Science writing or representations are used by students		
Sub Total		
Content		
l) Content is aligned to the Core		
m) Science Misconceptions were appropriately addressed		
Sub Total		

- OR -

Instructional time spent doing activities not described above. These activities can best be described as (circle all the apply): Lecture, Worksheet, Reading Science, Activity for Activity Sake, and/or Other (Describe)

Explanation of Descriptors

Section II Instruction

a) Select 3 important science questions the teacher poses. If a new question is posed that appears to be more central to the leaning use it in lieu of the lower scoring question. For each count the number of student responses up to 5. Calculate the average of the three for this score. Teacher's use of comments to extend the conversation should not be included as one of the questions, but responses to the extension should be counted in the three.

b) The number of relevant teacher prompts used to extend the discussion specific to a single question/topic. Questions used to extend the thinking or adjust thinking should be included for an individual ratio. The total number up to five is the score.

c) The teacher uses examples or analogies to clarify science concepts or principles. The examples and analogies must be connected to the science concept being taught and help clarify the understanding of science. **The total number up to five is the score**

d) Did the teacher support inter-student discussion? Follow the teacher in the classroom or during the discussion to see if the teacher engages the student in giving and receiving of ideas, information and/or discussion from other students. Rate the effectiveness of inter-student discussion. Did students discuss the science concepts being taught, etc. Add up the specific number of time the **teacher** promotes student discussion. **The total number up to five is the score**

e) Teacher uses accurate science language throughout the science activity. When errors are made by students, he/she effectively supports the students in using accurate science language. **One point for each accurate use of a science term up to 5 points**. If misuse of science terms is present, score should not exceed one.

f) Teacher directed students in science investigations – one point (maximum of 5) should be included for each of the following elements of science investigations: * formulate question, * identify * observation, *formulate a testable question, * identify variable as independent, dependent or control, *record data, *make inferences, *discuss limitations of findings, *analyze data (Elements found in ILO #1 in the Core).

g) Students move into investigations and engage in science skills and process - one point (maximum of 5) should be included for each of the following elements of science investigations: * formulate question, * identify * observation, *formulate a testable question, * identify variable as independent, dependent or control, *record data, *make inferences, *discuss limitations of findings, *analyze data (Elements found in ILO #1 in the Core). If investigations are whole class, maximum score of 3.

h) Students' understanding of the investigations was assessed using formal or informal assessments. Teacher uses questions to individuals and the group to clarify and to assess student learning.

i) Models were used in instruction by teacher to add to students understand of science concepts or principles. The models are explained by the teacher only = 1, the models are explained by the student only 3, the models are explained by the student and the teacher uses student explanation to extend the learning of others in the classroom. Scores of 2 and 4 are used when the criteria 3 or 5 are partially met.

j) Students explained the models developed and teacher used the explanations to assess student knowledge. Scores of 1 for student explanations that are inaccurate and the teacher does not use the opportunity for learning, 3 when the student explanation is accurate and the teacher accepts it without discussion to extend learning through talk and argument, 4 when the teacher requires explanations of the model that are evidenced based, 5 when the teacher or students requires evidence and the model is used to extend the science thinking to big ideas and principles.

k) Students used writing to represent their understanding of science ideas or observations they make. Diagram, graphs, charts, recording observations in notebook, etc. If this is done, 2, if multiple representations are used 3, if the students focus the attention on understanding the phenomena 4, and if the writing and diagrams provide insight into understanding the science concepts and are used to clarify concepts 5.

l) Content aligned to the Core –

- 0 – Instructional objectives are not stated
- 1 – Instructional objectives are stated but not aligned to the Core.
- 2 - Instruction targets process skills or ILOs only
- 3 – A clear objective is identify, but the instruction in not well aligned to the objective
- 4 – A clear objectives but no ILOs.
- 5- A clear objective is identify for instruction and the instruction target this objective and the indicator supporting the objective

m) During the instructions student misconceptions were addressed appropriately. 0-5 scale,

- 0- misconceptions were created by teacher or perpetuated
- 1- misconceptions recognized by the teacher but not addressed

- 2- student states a misconception, but the teacher does not recognize the misconception.
- 3- misconceptions were noted and teacher indicated the nature of these misconceptions
- 4- teacher corrected misconceptions during the lesson
- 5- teacher clarified the misconceptions and provided students with insight through discussion of how to understand the concept.

*PESTL Science Classroom Observation Protocol – Copyrights are exclusive property of Essential Teaching and Learning, LLC.

Appendix C

Capsule Rating of the Quality of the Lesson

Capsule Rating of the Quality of the Lesson

In this final rating of the lesson, consider all available information about the lesson, its context and the teacher's purpose, and your own judgment of the relative importance of the ratings you have made. Select the capsule description that best characterizes the lesson you observed. Keep in mind that this rating is not intended to be an average of all the previous ratings, but should encapsulate your overall assessment of the quality and likely impact of the lesson.

O Level 1: Ineffective Instruction.

There is little or no evidence of student thinking or engagement with important ideas of mathematics/science. Instruction is highly unlikely to enhance students' understanding of the discipline or to develop their capacity to successfully "do" mathematics/science. Lesson was characterized by either (select one below):

> ° Passive "Learning" Instruction is pedantic and uninspiring. Students are passive recipients of information from the teacher or textbook; material is presented in a way that is inaccessible to many of the students.
>
> ° Activity for Activity's Sake Students are involved in hands-on activities or other individual or group work, but it appears to be activity for activity's sake. Lesson lacks a clear sense of purpose and/or a clear link to conceptual development.

O Level 2: Elements of Effective Instruction

Instruction contains some elements of effective practice, but there are serious problems in the design, implementation, content, and/or appropriateness for many students in the class. For example, the content may lack importance and/or appropriateness; instruction may not successfully address the difficulties that many students are experiencing, etc. Overall, the lesson is very limited in its likelihood to enhance students' understanding of the discipline or to develop their capacity to successfully "do" mathematics/science.

O Level 3: Beginning Stages of Effective Instruction.

Instruction is purposeful and characterized by quite a few elements of effective practice. Students are, at times, engaged in meaningful work, but there are weaknesses, ranging from substantial to fairly minor, in the design, implementation, or content of instruction. For example, the teacher may short-circuit a planned exploration by telling students what they "should have found"; instruction may not adequately address the needs of a number of students; or the classroom culture may limit the accessibility or effectiveness of the lesson. Overall, the lesson is somewhat limited in its likelihood to enhance students' understanding of the discipline or to develop their capacity to successfully "do" mathematics/science.

O Level 4: Accomplished, Effective Instruction

Instruction is purposeful and engaging for most students. Students actively participate in meaningful work (e.g., investigations, teacher presentations, discussions with each other or the teacher, reading). The lesson is well-designed and the teacher implements it well, but adaptation of content or pedagogy in response to student needs and interests is limited. Instruction is quite likely to enhance most students' understanding of the discipline and to develop their capacity to successfully "do" mathematics/science.

O Level 5: Exemplary Instruction

Instruction is purposeful and all students are highly engaged most or all of the time in meaningful work (e.g., investigation, teacher presentations, discussions with each other or the teacher, reading). The lesson is well-designed and artfully implemented, with flexibility and responsiveness to students' needs and interests. Instruction is highly likely to enhance most students' understanding of the discipline and to develop their capacity to successfully "do" mathematics/science.

Horizon Research, Inc. Inside the Classroom: Observation and Analytic Protocol – Page 9 11/30/00

CURRICULUM VITAE

NANCY E. HAUCK

Employment Address
Department of Education
Dixie State College of Utah
225 South 700 East
St. George, UT 84770
Office Phone: (435) 652-7844

Home Address
886 E 970 S Circle
St. George, UT 84790
Home Phone: (435) 673-8855
Cell Phone: (435) 862-8987
Email: hauck@dixie.edu

Education

- 2008-2012 Ph.D. Curriculum & Instruction with Emphasis in Elementary Science Education, Utah State University, Logan, Utah
- 1992 M.Ed. Human Development & Literacy, Pacific University, Forrest Grove, Oregon.
- 1988 B.S. Elementary Education & Family Consumer Science Secondary Education, Southern Utah University, Cedar City, Utah.
- 1986 A.S. Dixie State College, St. George, Utah.

Teaching Experience in Higher Education

1993-Present Dixie State College (DSC), St. George, Utah

(2003-present) Full-time Tenured Assistant Professor

(1993-2002) Part-time Advisor, Field Placement Director, and Adjunct Instructor in Education & Family Studies Department

- Teaching Responsibilities: (currently) 15 semester credit hours, Methods of Elementary Science, Methods of Health & Movement, Fine Arts in Education, Principles of Early Childhood Education, Foundations of Multicultural Education
- Field Placement Director Responsibilities: (2003-2009) Arrange all education field placements
- Education Advisor Responsibilities: (1993-2002) Advise students, application profiles for education program candidates, counsel students as needed

- Classes Taught: (past & present) Methods of Elementary Science, Methods of Health & Movement, Fine Arts in Education, Principles of Early Childhood Education, Foundations of Multicultural Education, Curriculum Design & Assessment, Foundations in Education, Children's Literature, Introduction to Exceptional Learners, Human Development Through the Lifespan, Family Relationships

1995-2002 Southern Utah University (SUU), Cedar City, Utah
Education Adjunct Instructor and Advisor

- Education Advisor Responsibilities: Advise education majors at DSC/SUU University Center
- Classes Taught: Teaching Creative Arts in Education, and Foundations in Education

1990-1992 Pacific University, Forest Grove, Oregon
Professor's Assistant & Graduate Studies

- Graduate Studies: Research assistant for biological human development at the In-vitro Fertilization Experimental Laboratory, Oregon Regional Primate Center, Portland, Oregon
- Professor's Assistant: Lecture assistant for Dr. Patrick Conway in Personality Theory and Developmental Psychology courses
- Graduate Studies: Teaching reading to elementary students in the Pacific University Reading Lab

Teaching Experience in Elementary & Early Childhood Education

2010-present Washington County Schools, St. George, Utah
Astronomy Education Programs Elementary and Intermediate Schools

- Astronomy program coordinator in partnership with DSC Science Department

1993-2002 East Elementary, St. George, Utah
Science, Culture & Art Specialist Volunteer K-5

1998-2002 Tonaquint Nature Center, Parks & Recreation, St. George, Utah
Teacher for Nature Education K-6

1993-2002 Montessori Children's House, St. George, Utah
Science, Culture & Art Specialist PK-2

1992-1998 Joy School, St. George, Utah
Teacher

1989-1990 Children's World, Portland, Oregon
Teacher

Experience in Science & Science Education

1993-Present Dixie State College, St. George, Utah
Full-time Tenured Assistant Professor (2003-present)

- Teaching Responsibilities: Methods of Elementary Science, Methods of Health & Movement

2010-present Washington County Schools, St. George, Utah
Astronomy Education Program Elementary and Intermediate Schools

- Astronomy program coordinator in partnership with DSC Science Department

1993-2002 East Elementary St. George, Utah
Science Specialist K-5

1998-2002 Tonaquint Nature Center, Parks & Recreation, St. George, Utah
Teacher for Nature Education K-6

1993-2002 Montessori Children's House, St. George, Utah
Science Specialist PK-2

1990-1992 Oregon Regional Primate Center, Portland, Oregon
Graduate Research Assistant in In-vitro Fertilization Experimental Laboratory

Publications & Presentations

Effects of Sustained Teacher Professional Development on the Classroom Science Instruction of Elementary School Teachers, Poster Presentation at ASTE (Association of Science Teacher Educators) 2012 International Conference, January 4-7, Tampa FL.

Teaching Children Science Through Nature, Presentation at the 2011 Copenhaver Institute, June 20-24, Roanoke College, Roanoke, VA.

Competent & Caring Mentors, Presentation at the 2011 Copenhaver Institute, June 20-24, Roanoke College, Roanoke, VA.

Association Memberships & Committees

2011-present	ASTE (Association of Science Teacher Educators)
2010-present	NSTA (National Science Teachers Association)
2007-present	UATE (Utah Association of Teacher Educators)
2009-present	Dixie State College NWCCU Accreditation Steering Board
2004-present	Dixie State College Faculty Senate
2007-2009	Strategic Enrollment Management Council
2000-2009	Dixie State College Alumni Board
2006-2007	Professional Educators Coordinator Committee
2000-2002	Dixie State College Education Program Conceptual Committee