2014

# Investigating the effectiveness of a professional development program for secondary physics teachers

Meghan Reynolds
*University of Northern Iowa*

# INVESTIGATING THE EFFECTIVENESS OF A PROFESSIONAL DEVELOPMENT PROGRAM FOR SECONDARY PHYSICS TEACHERS

An Abstract of a Thesis

Submitted

in Partial Fulfillment

of the Requirements for the Degree

Master of Arts

Meghan Reynolds

University of Northern Iowa

May 2014

ABSTRACT

National studies have exposed a shortage of highly-qualified high school physics teachers, with over 25% of physics teaching vacancies in Iowa being very difficult to fill or unable to be filled. In an effort to improve the nation's education system, there have been broad changes in the framework of what constitutes high-quality teaching which supports a trend away from traditional, lecture-based teaching methods and towards interactive, student-centered methods. This study analyzed the impact of a professional development program that aimed to prepare more high-quality high school physics teachers by building physics content knowledge through the use of reformed teaching techniques that could then be applied in their classrooms.

The Iowa Physics Teacher Instruction and Resources (IPTIR) program was a three-year professional development program with a total of 35 high school science teacher participants. Program staff administered conceptual and pedagogical evaluation to teachers, and collected conceptual assessment data from their students. Analysis of this data provided insight into the program's effectiveness as well as implications for future professional development programs.

The IPTIR program enabled 20 out-of-field high school physics teachers to obtain a State of Iowa physics teaching endorsement, and improved the content knowledge of the teachers and their students through the use of interactive engagement techniques such as PRISMS PLUS learning cycles and Modeling Instruction. The results of this study reveal the effectiveness of programs such as IPTIR, and emphasize a need for further similar programs to produce more quality high school physics teachers.

# INVESTIGATING THE EFFECTIVENESS OF A PROFESSIONAL DEVELOPMENT PROGRAM FOR SECONDARY PHYSICS TEACHERS

A Thesis

Submitted

in Partial Fulfillment

of the Requirements for the Degree

Master of Arts

Meghan Reynolds

University of Northern Iowa

May 2014

This Study by: Meghan Reynolds

Entitled: Investigating the Effectiveness of a Professional Development Program for Secondary Physics Teachers

has been approved as meeting the thesis requirement for the

Degree of Master of Arts in Science Education

| | |
|---|---|
| Date | Dr. Lawrence Escalada, Chair, Thesis Committee |
| Date | Dr. Jeff Morgan, Thesis Committee Member |
| Date | Dr. Jody Stone, Thesis Committee Member |
| Date | Dr. Michael J. Licari, Dean, Graduate College |

# ACKNOWLEDGEMENTS

The process of writing this thesis has been a tremendous learning experience for me, and it's been a privilege to undertake this body of work. The final product would not have been possible without the guidance from my thesis advisor, committee members, and the support from my family and friends.

It is with immense gratitude that I acknowledge Dr. Lawrence Escalada, my graduate thesis advisor, for welcoming me into the IPTIR project and providing continual feedback and insight. His calm, patient guidance kept me on track and motivated me to always put forth my best effort. In addition to Dr. Escalada, I would like to thank the additional members of my thesis committee, Dr. Jody Stone and Dr. Jeff Morgan, for their contributions to the data analysis and writing process which greatly contributed to the quality of this work.

This thesis would not have been possible without Dr. Escalada, Dr. Stone, Dr. Morgan, and the additional members of the IPTIR faculty, Ms. Karen Couch-Breitbach and Mr. Les Burns. I give great thanks for their expert execution of the IPTIR program and for collecting the wealth of data I was provided for this study.

I would like to extend sincere thanks to my family for their unwavering support. To my siblings, you've all given your little sister a role model, each with your own strengths for me to look up to. And to my parents, every time I use questions to scaffold a student's learning, or use horribly dorky (yet effective!) analogies, these I can attribute to you and make me the teacher I am. Thank you for raising me to value learning and education, I wouldn't be the person I am without you.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

FIGURE PAGE

# CHAPTER 1

# INTRODUCTION

Today's technological era is one of exponential advancements in medicine, engineering and technology with seemingly no limit to what can be discovered and learned. This progress has importance both economically for the nation as well as for the 'greater good' of human knowledge and advancement. As Karl Fisch and Scott McCloud's 2012 video "Did you Know?" said:

> Did you Know… It is estimated that a week's worth of the New York Times contains more information than a person was likely to come across in a lifetime in the 18th century (3:09)… We are preparing students for jobs that don't even exist, using technologies that haven't been invented, to solve problems we don't even know are problems yet (0:47).

A constant influx of new information and technologies implies that education must prepare students to be scientifically and technologically literate. Because it is unknown what jobs will exist in ten years, student education can't just focus on *how* to do things, rather students need to know how to *figure it out*. For many students, these skills are built in high school science classes. Physics at its core is about figuring out how and why things happen, therefore physics classes hold great potential for building the science reasoning skills that will prepare students for the 21st century workforce

Physics lies at the foundation of technology and engineering, a field that is becoming increasingly more dominant in the classroom. The nationally developed Next Generation Science Standards include the previously overlooked field of engineering practices (Next Generation Science Standards, 2013), emphasizing the need for a focus in the areas of engineering and its physics underpinnings. Today's high school physics teachers must be able to provide quality, effective and meaningful physics education in order to build the foundation for the nation's future workforce.

In February 2010, The Task Force on Teacher Education in Physics (T-TEP) performed a national investigation into the number of quality of U.S. physics teacher preparation programs, and the research, policy and funding implications that foster effective physics teacher education (Task Force on Teacher Education in Physics, 2012). The Task Force reported that with only a few exceptions, physics teacher preparation in the U.S. is inadequate, inefficient, and unprepared to meet the needs of the 21st century physics student. The report concluded that significant changes need to occur on the state and national level in physics teacher education in order for every student to have the opportunity to learn physics from a knowledgeable, highly qualified physics teacher.

These findings beg the question, "what makes a quality physics teacher?" Physics builds upon foundational concepts as found in the Iowa Core, including knowledge of motion and forces, energy and its interactions, magnetism, and electricity (Iowa Department of Education, 2012a). A deep understanding of these concepts is a basic component of a quality physics teacher.  However, quality teaching goes beyond mastery of content; a quality physics teacher also has the expertise and ability to teach these concepts clearly. Quality teaching includes proficiency in scientific inquiry, such as engaging students with questions and creating scenarios where students can formulate explanations based on evidence, as emphasized in the Iowa Core (Iowa Department of Education, 2012a).

## Statement of Problem

### High School Physics Teacher Preparation

Deficiencies persist in physics teachers' ability to instill deep understanding, and one reason for this deficiency is often a gap between pedagogy and content knowledge in today's physics teacher preparation. Much of today's teacher education focuses on teaching methodology and theory, disconnected from the pre-service teacher's content area (Loughran, Mulhall & Berry, 2008). Emphasis in teacher education is placed on classroom management, lesson planning, assignments, and time allocation (Shulman, 1994). While these are valuable skills to learn, emphasis on these aspects of teacher education without equal emphasis on content-specific teaching pedagogy can lead to the inability to effectively teach difficult concepts within the teacher's content area.

In a 2008-2009 the American Institute of Physics (AIP; 2012) surveyed approximately 3,600 public and private high schools across the U.S. gathering data on physics teacher background, preparation and course characteristics. The survey found that approximately 24% of high school physics teachers have a major in physics or physics education. The Taskforce on Teacher Education in Physics (T-TEP) concluded that over 90% of the teachers that *do* receive an undergraduate degree in physics education were educated by programs that do not provide coherent content and pedagogical preparation (T-TEP, 2012). As a result a majority of physics teachers enter the classroom unprepared and are faced with a "trial and error" approach to learning pedagogical skills.

Physics Teacher Shortages

Since 1987, there has been a 28% nationwide increase in the number of teachers who taught at least one physics class (17,900 in 1987 to 23,000 in 2012), however growth in the number of physics teachers is much smaller than the growth of the number of physics students (Tesfaye & White, 2012). In 2012, the Iowa Department of Education (Iowa DoE) designated all science subjects Grades 5-12 as teacher shortage areas, as determined by the number of Class B licenses issued as well as the projected number of science teacher graduates for the year (Iowa DoE, 2012b). A class B teaching license is a conditional license that allows a teacher to teach outside his/her endorsement area while concurrently obtaining the necessary educational credits for the outside endorsement area.

In 2011-2012, the National Center for Education Statistics found that 25.5% of physical science teaching vacancies nationally were either very difficult to fill or not able to be filled (U.S. Department of Education, 2011). As a result, many physics teaching vacancies are filled by out-of-field physics teachers – those whose area of expertise is in other science subjects and who have yet to meet the physics course requirements for a physics teaching endorsement (Escalada & Moeller, 2006). These teachers are often short in the number of physics credit hours they need to meet the requirements for a physics teaching endorsement, and these requirements must be met within a certain period of time while teaching physics.

## Theoretical Framework

Physics education research focuses primarily on student understanding of physics concepts, how instruction alters these concepts, and the problem-solving techniques of students (Knight, 2004). These make up the facets of student learning, which are the specific strategies and pieces of knowledge students use to solve problems and answer questions. Facets can be correct, or can be a misconception or false observation, and are valuable for analyzing students' thinking and decision processes (Minstrell, 2000). To capture each of these facets of student learning, research involves not only physics but also psychology and cognitive science, providing a framework from which best educational practices can be determined. Today's physics education research considers not only what knowledge to instill in students but also the knowledge students bring to the classroom. This view of instruction aligns directly with the constructivist model of teaching, where the individual learning experience is not that of a passive onlooker, rather it is created from within the learner (Staver, 1998).

According to the constructivist model, the individual acquires scientific knowledge by constructing their own concepts, and this construction is a product of both the knowledge already possessed and the new information being received (McDermott, 1991). Prior knowledge, particularly of physics concepts, is often inconsistent if not incorrect, and these misconceptions are quite resilient to change (Knight, 2004). Misconceptions, also called naive ideas or alternate conceptions, are erroneous understandings of concepts and/or phenomena (Goris & Dyrenfurth, 2010). Misconceptions are based on students' prior knowledge and have been internally justified and reasoned, believed to be true by the student but in reality represent flawed understanding. Therefore it is essential to not just inform students of physics concepts, they must understand and be convinced of the behavior of physical phenomena in order to dispel their misconceptions. From a constructivist framework, physics education research has identified implications for instruction that provides the student with opportunities to construct this knowledge and models to conceptually represent and understand physics phenomena.

## Identification of Terminology

*Content knowledge* refers to knowledge of subject matter and the intricate ways in which it can be associated with different degrees of cognitive complexity (Shulman, 1987). As a teacher, sufficient subject knowledge is essential not only to be able to instruct and define basic concepts, but to also to guide students to discover *why* the concept is so.

*Pedagogical content knowledge* goes beyond the subject matter to the ability of a teacher to translate knowledge to students. An effective teacher uses analogies, explanations, demonstrations, and to represent content knowledge to a student. Without this knowledge, the teacher often misrepresents the information and can cause student misconceptions (Halim & Meerah, 2002). This can lead to frustrations on the part of both teacher and student, which can contribute to negative view points on the topics (Ucar, 2012). However, with sufficient pedagogical content knowledge a teacher can present the concepts of physics at a cognitively complex level, allowing the motivated student to strengthen his or her understanding of the physical world.

*Traditional teaching* refers to conventional instruction that involves the teacher directly providing students with necessary information, and as such is often referred to as *direct instruction*. Students are passive receivers of information and classrooms are *teacher centered* (Wells, Hestenes & Swackhamer, 1995).

*Reformed teaching* is defined by educational research into what teaching methods are more effective than others (Knight, 2004). The reformed teaching framework is based on student cognition, prior conceptions and a need to organize knowledge to achieve functional understanding. The role of the teacher is to facilitate the development of student knowledge and utilize interactive engagement techniques by questioning students, therefore classrooms are *student centered*.

*Interactive engagement* methods promote student conceptual understanding through the use of probing questions, student discussion and hands-on activities (Hake, 1998).

*Conceptual understanding* goes beyond the symbols, numbers, procedures and facts that define specific knowledge, to the connections between ideas and application of concepts. Conceptual understanding requires higher-order thinking as compared to memorizing facts, and therefore cannot be achieved by rote memorization. Conceptual knowledge is achieved through understanding of the underlying concepts, and the knowledge can be reconstructed and applied to various scenarios (Balka, Hull, & Harbin Miles, 2013).

The *Learning Cycle* is a structure of teaching and learning intended to guide students towards questions and a desire to understand the concept at hand. This is followed by methods that ideally allow the students to discover answers to their questions and then apply the new concept to new situations. The three stages of the learning cycle are commonly referred to as Exploration, Concept Development and Introduction, and Application. Other variations of the learning cycle exist including the 5E learning cycle consisting of Engagement, Exploration, Explanation, Elaboration, and Evaluation (Bybee et al., 2006).

*Out-of-field physics teachers* are those whose area of expertise is in other science subjects and have yet to meet the physics course requirements for a physics teaching endorsement.

*Facets* of student learning are the specific strategies and pieces of knowledge students use to solve problems and answer questions. Facets can be correct, or can be a misconception or false observation, and are valuable for analyzing students' thinking and decision processes (Minstrell, 2000).

*Misconceptions*, also called naive ideas or alternative conceptions, are erroneous understandings of concepts and/or phenomena. Misconceptions are based on students' prior knowledge which have been justified and reasoned. These are ideas formed from everyday experiences to create understanding believed to be true by the student, but in reality represent flawed understanding. To overcome misconceptions, students must encounter a conceptual change and new understanding, and often misconceptions are very resistant to change (Goris & Dyrenfurth, 2010).

*PRISMS PLUS*, developed with funding from the National Science Foundation in collaboration with master high school physics teachers, is a learning-cycle based physics curriculum containing 44 complete cycles in four unit books covering Force and Motion, Work and Energy, Waves and Optics and Electricity, and Modern Physics (Cooney, Escalada & Unruh, 2008).

*Modeling Instruction* is a modified learning cycle curriculum developed in response to persistent weaknesses in traditional, lecture-based teaching methods. Modeling Instruction uses engaging, insightful activities to give students opportunities to build conceptual models to aid in the understanding of physical phenomena (Jackson, Dukerich, & Hestenes, 2008).

*Socratic Questioning* is often associated with Modeling Instruction, and is the practice of using question prompts to promote reflective and critical student thinking (Wenning, 2005). The questions, which can be posed by both teacher and classmates, hold students accountable for their own learning by exposing thinking processes and conceptual understanding (or lack thereof).

*Whiteboarding* is a Formative Assessment Classroom Technique (FACT) often associated with Modeling Instruction, and usually involves small groups of students collaborating to present explanations, graphs, tables, and/or diagrams on a dry erase board. The boards are generally large (24"x32"), allowing both teacher and students to quickly discern a group's explanation and reasoning (Keeley, 2008). Whiteboarding encourages an environment where students generate their own ideas and solutions. Students explain their findings to a question previously posed to the class, followed by dialogue involving other classmates and the teacher, often involving Socratic questioning (Wenning, 2005).

The *Force Concept Inventory* (FCI) is a conceptual assessment which tests comprehension of basic Newtonian concepts. The test is comprised of 30 multiple-choice questions covering six major Newtonian concepts: (1) Kinematics, (2) Newton's First, (3) Second, and (4) Third Laws, (5) Superposition, and (6) Kinds of Force (Hestenes, Wells & Swackhamer, 1992). The test is structured to require a choice between Newtonian concepts and common alternative misconceptions (Huffman & Heller, 1995).

The *Test of Understanding Graphs in Kinematics* (TUG-K) is a multiple-choice conceptual assessment designed to uncover common difficulties and misconceptions students often have interpreting kinematic graphs (Beichner, 1994).

The *Classroom Test of Scientific Reasoning* (Lawson Test) assesses deductive and inductive reasoning, proportionality and probability reasoning, and other processes of scientific thinking across various realms of science (Lawson, 1978).

The *Reformed Teaching Observation Protocol* (RTOP) is a tool used to quantitatively analyze reformed teaching practices. It is a 100 point Likert-scale instrument broken down into five 20-point categories: Lesson Design and Implementation, Propositional Knowledge (Content), Procedural Knowledge (Inquiry), Communicative Interactions, and Student/Teacher Relationships (Sawada et al., 2000).

*Iowa Physics Teacher Instruction and Resources* (IPTIR), a three-year professional development program, was developed to prepare more high-quality secondary physics and physical science teachers for Iowa's schools. Utilizing Modeling Instruction, PRISMS PLUS learning cycles and other interactive engagement methods, the program developed participants' physics pedagogical content knowledge using methods participants could then apply in their classrooms.

*Average normalized gain* is the average increase in scores divided by the average increase that would have resulted if all students had perfect post-test scores, and is used in this study to estimate improvement or average teaching effectiveness (Coletta & Phillips, 2005).

*Average normalized change* is used when a decrease is encountered from pre-test to post-test, and is the average decrease in scores divided by the maximum possible decrease, which is also represented by the pre-test score (Marx & Cummings, 2007).

## Previous Studies

Extensive research has been performed to determine why deficiencies persist in today's physics teachers, and many studies identify the root of the problem to be physics teacher education. To teach a concept well a teacher must possess deep understanding of the concept, which is often not being achieved in physics teacher education. Without strong conceptual understanding, a teacher's knowledge is often fragmented, compartmentalized, and poorly organized so that when instructing, this knowledge is difficult to access (Loughran et al., 2008). Studies identify specific aspects of physics that pre-service physics teachers have increased difficulty learning (Sahin & Yağbasan, 2012) and how these difficulties result in a lack of content knowledge which translates to misinformed instruction (Halim & Meerah, 2002). Mathematical competency can be a large factor in understanding physics concepts and difficulties in math likely parallel difficulties in physics (Sahin & Yağbasan, 2012).

Studies report that misinformed instruction can lead to student misconceptions and a fragmented understanding of physics concepts (Ucar, 2012), and that this chain of events can be broken by improving both the content knowledge and pedagogy of both in-service and pre-service physics teachers (van Driel, Verloop & de Vos, 1998). Previous professional development programs report promoting deeper scientific understanding and inducing an overall change in scientific teaching (Yerrick, Parke, & Nugent, 1997) and the most effective professional development utilizes methods to instill content knowledge that are in turn directly applicable in the teacher's classroom.

## Deficiencies in Previous Studies

Previous studies clearly outline the need to improve secondary physics teacher education and provide professional development aimed at expanding physics teacher content knowledge and pedagogy. Only a limited number of studies have assessed the effectiveness of these programs and the impact the expansion of teacher knowledge and pedagogy has on student learning. Previous studies also are limited in that they focus on master physics teachers at the college and high school level but not on novice physics teachers over a period of time. This study aims to determine overall effectiveness of research-based instructional methods used to guide novice high school physics teachers to reformed classroom practices, as well as take a closer look at student impact and determine aspects of the program that can give insight to future professional development programs.

Purpose and Significance of Study

The purpose of this study is to both quantitatively and qualitatively analyze changes among participants in a physics teacher professional development program aimed at improving physics teacher pedagogical content knowledge using research-based instructional methods. Evaluation includes analysis of participant content knowledge and pedagogical methods as well as the transfer of these changes into the classroom based on classroom practice as well as student performance. Teacher feedback and conceptual understanding of both the teacher participants and their respective students throughout the program illustrates the program's initial, intermediate, and long-term outcomes (Escalada, Morgan & Stone, 2010).

By evaluating the effectiveness of this professional development method, the study will in turn provide significant guidance to other physics teacher professional development programs. An increase in the number of effective physics teacher programs nationwide can lead to an increase in the number and quality of physics teachers with endorsements in physics, addressing the current shortage of qualified high school physics teachers. Additionally, the effective professional development methods may lead to similar education initiatives in other content areas, leading to improvements in overall student performance on standardized tests as well as an increase in comprehensive knowledge (Escalada et al., 2010).

# CHAPTER 2

# REVIEW OF LITERATURE

## Pre-Service Physics Teacher Education

In response to the nationwide shortage of physics teachers and lack of effective physics instruction, the Task Force on Teacher Education in Physics (T-TEP) was formed with the mission of determining the state of physics-teacher preparation and to provide recommendations for higher quality physics-teacher preparation (T-TEP, 2012). T-TEP surveyed all 758 U.S. collegiate physics departments, and within the high-producing departments (two or more physics education graduates a year) faculty members were interviewed to supplement the survey. In order to recognize the leaders in physics-teacher preparation T-TEP then conducted site visits to the most promising institutions exhibiting high-quality program. The Task Force found that these high-quality programs were rare, and that the great majority of physics teacher preparation programs were ineffective at producing teachers prepared to meet the needs of the nation's students.

One significant finding of the T-TEP was that with little exception, pre-service physics teachers receive education on their specific content as well as teaching methods and theory, but little in the way of content specific teaching methodology. Rather, educational coursework addressed state certification requirements and did nothing to develop physics pedagogy (T-TEP, 2012).

Ideally, teachers' understanding of physics concepts is developed during undergraduate education, and views on the nature of science are developed even prior to this stage. Loughren et al. (2008) investigated beyond the development of student pedagogical content knowledge into how content knowledge and the understanding of the nature of science influences pre-service science teacher's views about teaching science. They found that this simple awareness of content knowledge induced a shift in pre-service teacher thinking towards aligning content with pedagogy. Introductory physics courses taken at the undergraduate level coupled with previous exposure to science create the foundation from which deep understanding can be built upon and then accessed when teaching.

One study found that the most difficult concepts for pre-service physics teachers are magnetism and electromagnetic waves. The researchers projected this difficulty was from a lack of mathematical background knowledge needed to understand the concept (Sahin & Yağbasan, 2012). Ucar (2012) reported that teachers' negative attitudes towards difficult concepts often stem from frustrations because they have not been instructed using inquiry-based teaching methods during their pre-service training. After examining a science training program for pre-service teachers, Ucar found that exposure to scientists and scientific environments promoted more positive views and therefore more effective teaching strategies.

Van Driel, Jong and Verloop (2002) examined pedagogical content knowledge in pre-service chemistry teachers and their ability to switch from macro-level observations, such as a chemical reaction, to atomic level thinking. This study can be applied to physics as well, such as the macro-level properties of an electrical circuit and micro-level understanding of the interactions between the electrons within the circuit. The teachers' ability to switch between these "modes" of thinking was improved by increasing their content knowledge, which was done via workshops and mentors. The importance of pre-service teachers' understanding was made evident in a study that concluded that pre-service teachers with little understanding were more likely to be unaware of student misconceptions (Halim & Meerah, 2002). These trainee teachers were impeded by their poor content knowledge, and when attempting to form analogies for instruction the teachers often instead created further student misconceptions.

## Current Status of Secondary Physics Teachers

Recent studies capture a snapshot of the nation's secondary physics teachers in terms of demographics, preparation, pedagogy and views on teaching. In order to develop effective professional development an understanding of the target audience must first be established.

### Academic Preparation

A 2008-2009 study of about 3,600 public and private high schools across the U.S. by the AIP found that only 26% of high school physics teachers have a degree in physics or physics education (AIP Statistical Research Center, 2012). In a 2012 Horizon Research study of almost 650 high public and private high school data indicated that only 20% of the physics teachers had a degree in physics, and 14% having never taking a college course in physics (Banilower et al., 2013).

When Horizon Research surveyed whether the teachers consider themselves to be "very well prepared" to teach various topics, results indicated a range of self-reported preparedness across physics topics (Banilower et al., 2013). Of the teachers surveyed, 80% felt prepared in areas of forces and motion, 54% felt prepared in electricity and magnetism, and 17% felt prepared in engineering related topics. In a parallel survey by AIP (2012), 98% of those surveyed reported feeling adequately or well prepared to teach physics, which while this can't be directly compared to the reported "very well prepared" it gives a window into the range of teachers that perceive themselves as only adequately prepared to teach physics.

Beliefs About Teaching

The AIP and Horizon Research studies also surveyed high school physics teachers on their classroom activities and pedagogy. In 2008 AIP reported that 95% of teachers surveyed used traditional lecture at some point, and 34% used this most often. In contrast 65% reported using activity-based guided-inquiry, and 10% used this activity most often. In a similar vein, the 2012 Horizon Research survey revealed that 64% of those surveyed felt that definitions should precede the instruction on a science idea, and 45% of teachers reported that lab activities should be used to reinforce ideas already learned (Banilower et al., 2013). Horizon Research also reported that 92% indicated science instruction should focus on depth not breadth, and 92% that class should provide opportunities for student thinking and reasoning.

## Professional Development for Physics Teachers

Research has established that graduates of the majority of physics education programs receive little in the way of training in physics pedagogy (T-TEP, 2012). In addition to pedagogically ill-prepared physics education graduates, because of the high need for physics teachers many come from other educational, science and/or occupational disciplines. Professional development is the engagement teachers, especially those coming from alternative fields, can use to develop the necessary content and pedagogy skills to be high-quality, effective physics teachers. Additionally, as a teacher's college career becomes more distant in the past the importance grows to remain in touch with the progress and change in science education (Bucher, 2009).

When analyzing and developing teacher content knowledge, it is important to distinguish pedagogical content knowledge from both general pedagogical skills and knowledge of the subject matter (van Driel et al., 1998). When developing content knowledge, meaningful, content-rich professional development improves teachers' content preparation which is a powerful tool for effective teaching (Supovitz & Turner, 2000).

<u>Previous Professional Development Programs</u>

Professional development comes in many forms and at many levels, from individual teacher development to state-wide initiatives. The focus of the professional development may vary from developing specific content knowledge to developing teaching pedagogy techniques, as well as combinations of both content and pedagogy. Many teachers may pursue continuing education in specific content areas and/or pedagogy while working towards an advanced degree. Workshops and institutes provide exposure to new methodology and provide networking and mentoring experiences (American Association of Physics Teachers, 2009). Research or work experience gives teachers tools to incorporate real-world applicability into their classrooms.

The American Association of Physics Teachers (AAPT) in conjunction with the Physics Teaching Resource Agents (PTRA) introduced both the Urban PTRA program

(2000-2003) and the Rural PTRA program (2003-2008). Urban PTRA focused on high school physics teachers in large, urban school districts and consisted of weekend workshops on segmented curriculum topics with no required attendance commitment, and did not administer assessments (Matsler, 2010). Rural PTRA focused on secondary physics and physical science teachers with week-long summer institutes and follow-up sessions during the academic year for three years. The project curricula were coherent and modeled off of research-supported best practices, and administered assessments to both participants and students, thus discussion here will focus on the Rural PTRA.

The Rural PTRA program spanned across 11 U.S. universities all of which had trained PTRA leaders, and each site could invite up to 25 teachers (Matsler, 2010). Over the three years the program goal was to provide at least 108 instructional hours, in accordance with the No Child Left Behind act requirements for professional development, and to have the greatest impact on the participants and consequently their students. Participant impact included slight increases in the teachers' physics content understanding and significant increases in confidence level in both physics content and pedagogy of the teacher participants. Self-reported surveys indicated changes in classroom practices including incorporation of technology and shifts from teacher-centered to more student-centered classrooms featuring active student engagement. Broader impacts of the program included changes in participant classroom practices resulting in documented increases in student achievement, increased student interest in science, as well as systemic teacher preparation reform at the university level (Matsler, 2010).

From 2002 to 2003, the University of Northern Iowa (UNI) Physics and Science Education faculty conducted the UNI Physics Institute, a professional development program for secondary physics teachers (Escalada & Moeller, 2006). The 21 participants received interactive instruction in both physics content and pedagogical methods and completion of the program provided the necessary course requirements to attain a 7-12 physics teaching endorsement in the state of Iowa. The program consisted of two 4-week summer sessions and seminars during the academic year via communications network. The teaching methods focused on interactive engagement and learning cycle instruction.

Pre- and post-test data was collected from both the participants and their students to measure conceptual understanding. Analysis of results showed the impact of the participants' successful application of the program's interactive methods. The highest impact was from participants that scored at mastery levels on the conceptual assessments, indicating that the program was most appropriate for those with physics and/or mathematic backgrounds (Escalada & Moeller, 2006).

## Characteristics of Effective Professional Development

The ultimate goal of any professional development is effective implementation of the learning incurred during a program. A teacher may learn a great deal and exhibit gains in content knowledge and understanding of effective pedagogy, but if these skills aren't transferred into the classroom the professional development cannot be deemed effective.

When developing in-service teachers' content knowledge, the teachers must concurrently play the role of educator as well as student. The methods that are most effective at increasing teacher understanding should in turn be applicable in the teacher's classroom. Likewise, just as a goal for students is the ability to transfer knowledge to application, a goal for professional development is to transfer philosophy to practice. These goals can be accomplished using the same basic principles of utilizing active engagement, incorporating personal experience and knowledge, and addressing misconceptions to align conceptual understanding (Radford, 1998). The National Science Teachers Association (2013) recommends that science teachers experience inquiry as part of their professional development and that they themselves develop questioning strategies that will foster inquiry in their classrooms.

The Iowa Professional Development Model (IPDM) provides guidance and requirements for professional development programs, including rubrics which specify effective components within the various dimensions of a professional development program. The IPDM recommends that the design of any professional development should include research-based rationale for the strategies that are implemented, and that a program provides adequate time for teacher training that also extends through the school year (Iowa DoE, 2013). The IPDM advises that training be continuous with multiple opportunities to collaborate with other teachers to practice the newly learned content and pedagogy. Professional development programs funded with Title II grants in the state of Iowa are required to incorporate the Iowa Professional Development Model (Iowa DoE, 2012c).

Research on the distribution of professional development has shown that long term in-service teacher education distributed throughout the academic year is more effective than single events (Kennedy, 1998). Horizon Research, Inc. recommended a minimum of 80 hours in order to change classroom practice, and the AAPT/PTRA suggests professional development should be sustained for 80-100 hours and spread out over the course of several months or years to allow for implementation and modification of classroom practice (Matsler, 2010).

Multiple studies describe a necessary characteristic of a successful professional development is that science content and teaching methods be integrated and team taught by scientists, science educator, and classroom educators (Radford, 1998). Many teachers may experience a lack of confidence and/or cognitive dissonance while developing their skills, and instructor support can provide essential support over time throughout the classroom restructuring process.

## Reformed Physics Teaching Methods

In an effort to improve the nation's education system there has been a broad change in the framework of what constitutes high-quality teaching. In general this trend is moving away from traditional teaching methodologies involving the direct delivery of information to students, towards practices rooted in the constructivist philosophy that knowledge must be constructed by the learner. Many "buzz-words" circulate through the science education community to describe this constructivist philosophy, including reformed, inquiry-based, activity based, student-centered and interactive engagement teaching methods (National Institutes of Health, n.d.). Regardless of the label, this reformed model of teaching places the students in the classrooms as active participants in the learning process. Within science education, reformed teaching emphasizes the scientific process and encourages students to think logically and create experiments. Science is an investigative process, and reformed science teaching through inquiry is intended to reflect this process (National Science Teachers Association, 2013). The National Science Teachers Association recommends that science teachers utilize an inquiry-based curriculum that causes students to question, explore and raise questions, identifying the learning cycle as one possible effective strategy.

The Learning Cycle

The learning cycle structure of teaching and learning is structured to guide students towards questions and a desire to understand, followed by methods that ideally allow the students to discover answers to their questions. Research dating back to the early 20$^{th}$ Century identified the importance of allowing students the opportunity to discover relationships from their own experiences, followed by the explanation of ideas related to the student's initial discoveries, and finally a demonstration of understanding through application (Bybee et al., 2006). These components make up the core structure of a learning cycle, and though there is a variety of learning cycle versions with varying stages, they all fundamentally share these similar aspects.

In the initial stage of the three-stage learning cycle, Exploration, students explore a new concept or situation (Lawson, Abraham & Renner, 1989) and the activity should be as student-guided as possible. The new experiences ideally lead the students to questions, recognition of patterns, and/or observation of phenomena that typically goes against their current thinking. This stage can be an experiment where data is collected or an observation of an event, and done individually, in small groups or even as an entire class (Maier & Marek, 2006).

In the second stage, Explanation or Concept Development, data is analyzed and interpreted, and/or the teacher addresses the questions developed or patterns recognized by associating terms with concepts (Maier & Marek, 2006). The Explanation activity should relate directly to the Exploration stage (Lawson et al., 1989). Explanation may be done directly by the teacher or can be done via class discussion, with students identifying and/or answering their own questions as much as possible prior to teacher explanation.

The last phase (when using a three-stage cycle) is referred to as Elaboration, Application, or Concept Application, where students apply the new concept to a different context or scenario, and can also serve as assessment and/or a lead in to the next learning cycle (Maier & Marek, 2006). It is often essential for students to extend application past the initial scenario in order to move on from previous misconceptions or prevent learning from being isolated to only the initial example (Lawson et al., 1989).

The five-stage learning cycle, or 5E model, parallels the three-stage cycle at its core, but is expanded to include an engagement phase prior to exploration, and an evaluation stage following elaboration. The initial engagement phase is intended to connect the student's prior knowledge to the new concept, and the final evaluation phase assesses student understanding (Bybee et al., 2006)

An important aspect of the learning cycle is the development of science reasoning, a crucial skill in terms of transferring learned ideas to extended applications. Building upon Piaget's cognitive development theory, Robert Karplus applied previous research on student reasoning patterns specifically to science teaching (Karplus & Butts, 1977). Karplus recognized that an effective learning cycle allows teachers to identify, develop and assess student's reasoning patterns, guided by indicators explored by the teacher throughout the learning cycle.

Many science curricula effectively utilize the learning cycles, and within the field of physics one such example is PRISMS PLUS which consists of 44 complete learning cycles in four unit books covering Force and Motion, Work and Energy, Waves and Optics and Electricity, and Modern Physics (Cooney et al., 2008). The materials are the second generation of Physics Resources and Instructional Strategies for Motivating Students (PRISMS) which began in 1987 as a collection of approximately 130 engaging activities and respective teacher notes intended to span a year of high school physics curriculum. PRISMS PLUS employs an enhanced learning cycle structure, embracing the constructivist model of learning with activities focused on the development of scientific reasoning, scientific inquiry and conceptual understanding. The PRISMS PLUS curricula is used extensively in professional development programs and is commonly used as a supplemental resource in high school physics classrooms.

### Modeling Instruction

Much of the reformed teaching movement centers around preventing hollow learning, where students garner just enough information to regurgitate on a test. Even active classrooms can lack truly meaningful learning, with students completing the motions but without truly grasping understanding. In order to help students develop a coherent and systemic understanding of physics many teachers are adopting the modeling method of teaching physics (Wells et al., 1995). Wells et al. describe the modeling approach as one that surrounds a small number of basic models that describe patterns that appear throughout physics phenomena. The "model" is a conceptual representation of physics phenomena and the modeling method provides an alternative to the mindset that the student is supposed to find an answer as associated with traditional "plug and chug" problems. Rather the model is a solution and physics problems can be solved by adapting a known model. Modeling is intended to imitate the scientific process, focusing student attention on understanding rather than accumulating fragments of information (Wells et al., 1995).

Documented success supports the modeling method of instruction, which engages the learner in their understanding of the physical world by constructing and using scientific models (Hestenes, 1996). Instruction can be organized into modeling cycles to move the learner through the evaluation and application stages while breaking the models down into basic patterns (Wells et al., 1995). As students advance through the cycles they can develop explanations to account for their observations with data and other evidence to support these explanations that would otherwise be rooted in abstract conceptualizations (Windschitl, Thompson, & Braaten, 2008).

Modeling Instruction Workshops occur throughout the United States, often as three-week summer workshops often arranged via the American Modeling Teachers Association (AMTA). Modeling is implemented in these workshops as a means of improving content knowledge of the participants, and also gives the opportunity for the participants to develop interactive pedagogy skills. The AMTA provides extensive online resources, including a complete curriculum repository of Modeling units for physics as well as chemistry, physical science, and the developing biology Modeling units, all of which can be accessed by AMTA members (AMTA, n.d.). The units include teacher notes, worksheets, quizzes, tests, videos, and student readings to support the activities. The website also acts as a hub for external weblogs and forums, fostering an active community of collaboration among teachers utilizing Modeling instruction.

# CHAPTER 3

# METHODOLOGY

## Research Questions

The principle objective of almost any professional development program is to improve the teaching quality of its participants, and many aspects contribute to becoming a high-quality physics teacher. This study aims to determine if the structure and methods of the Iowa Physics Teacher Instruction and Resources (IPTIR) professional development program attained this principle objective by investigating the following:

1. What changes occur throughout the professional development program in regards to participant content knowledge?
2. Did the participants modify their teaching pedagogy in alignment with today's research-based instructional methods?
3. To what degree, if any, did changes in teacher content knowledge and/or pedagogy affect the classroom performance of the participants' students?

## The Iowa Physics Teacher Instruction and Resources (IPTIR) Program

The Iowa Physics Teacher Instruction and Resources (IPTIR) program was developed to prepare more high-quality high school physics teachers for Iowa schools and to improve the performance of their students by broadening teacher content knowledge and pedagogy to align with the latest research-based instructional methods and national/state science education initiatives. IPTIR provided a way for Iowa science teachers seeking a physics teaching endorsement to not only gain conceptual physics knowledge, but also acquire the skills for effective instruction of these concepts (Escalada et al., 2010).

Program Overview

The purpose of the IPTIR program was to address the critical shortage of qualified high school physics teachers by preparing more high-quality high school physics teachers for Iowa schools (Escalada et al., 2010). It was a three-year professional development program funded by the Board of Regents, State of Iowa and Title A of the No Child Left Behind Act, and a collaboration with University of Northern Iowa (UNI) Physics Education faculty, UNI Science Education faculty, master teachers, Area Education Agencies, and participating school districts.

From 2009 to 2012, the IPTIR program targeted a cohort of both existing high school physics teachers as well as secondary science teachers working towards the requirements necessary for the State of Iowa grades 5-12 physics teaching endorsement. The three-year program included two-week summer institutes each of the three summers as well as two Saturday meetings and two regional conferences during the academic years. Continual feedback and collaboration between the program faculty and participating teachers extended through the academic years. Participants were awarded three graduate credits for each summer and two graduate credits for each academic year of the program they completed. This allowed for participants to receive up to 15 graduate credits via the University of Northern Iowa, fulfilling the State of Iowa requirements for a grades 5-12 physics teaching endorsement.

The program focused on providing physics pedagogical content knowledge via modeling and other interactive engagement methods, including Physics Resources and Instructional Strategies for Motivating Students (PRISMS) PLUS (Cooney et al., 2008) and Modeling Instruction (Hestenes, 1987). The IPTIR program included two-week institutes each of the three summers, consisting of intense training on physics concepts. These summer institutes were followed by further professional development and support throughout the academic year.

The objectives of the IPTIR program, in alignment with the recommendations of national/state science education initiatives and physics education research, were as follows:

1. Prepare more high-quality secondary physics teachers that are knowledgeable in both physics content and pedagogy for Iowa schools;
2. Enhance the instructional practices of participating teachers consistent with national/state science education initiatives;
3. Improve the achievement of the participants' students in conceptual understanding of basic physics ideas and proficiency in science reasoning, and problem solving;
4. Expand the number of model high school physics classrooms that would provide positive learning environments for high school students as well as contribute to building the next generation of secondary science teaching majors as they complete their field teaching experiences and student teaching required for their professional education.

Program Evaluation Overview

To evaluate the effectiveness of the IPTIR program, data collection was performed using a mixed method technique with a variety of assessment measures. Figure 1 outlines the project logic model designed by the IPTIR program developers.

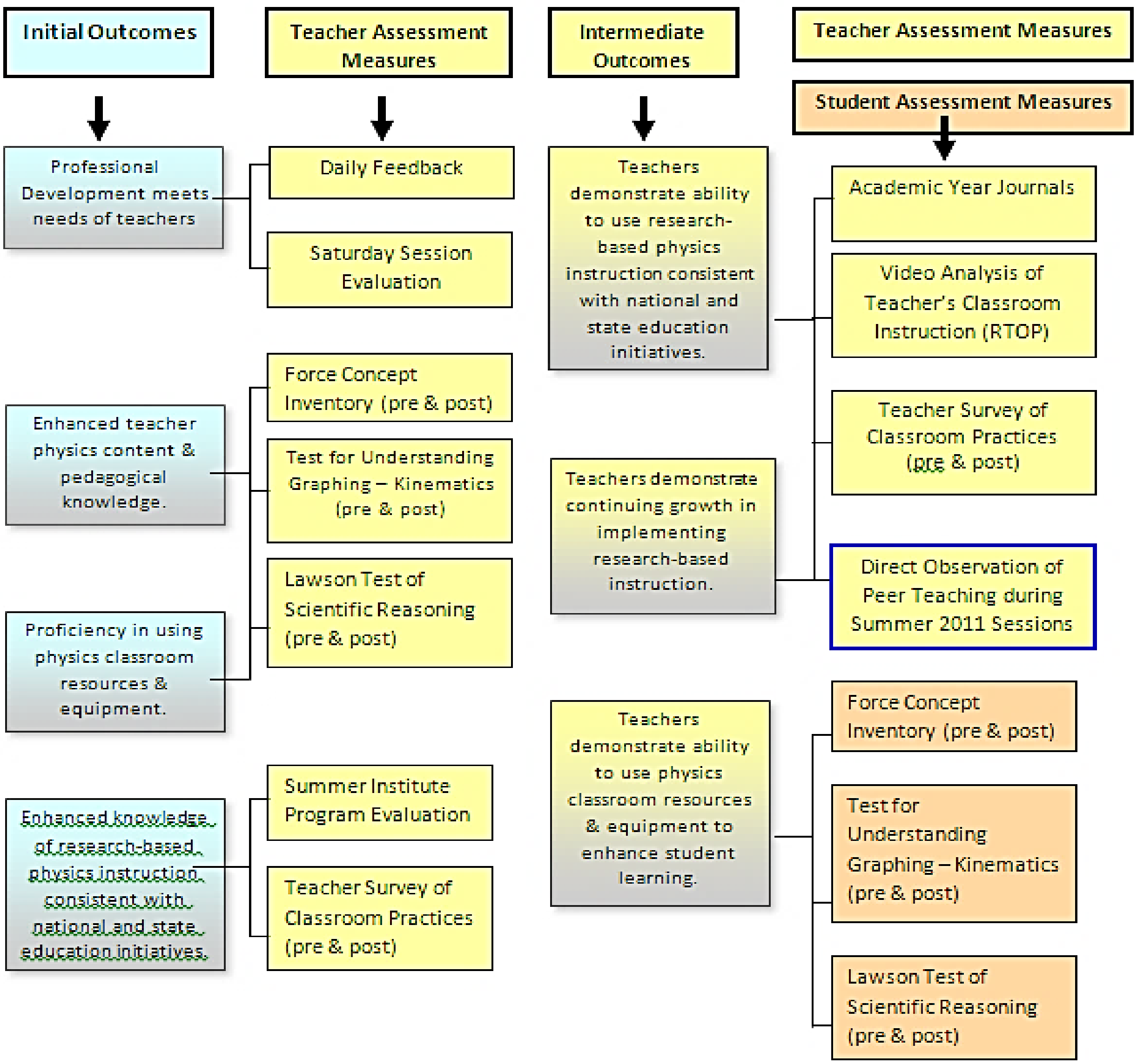



*Figure 1.* IPTIR Logic Model and Assessment Plan. Reprinted with permission from the IPTIR Project Annual Report by Larry Escalada, Jeff Morgan & Jody Stone, 2010.

The project model guided assessment throughout the length of the program and provided diagnostic, formative and summative assessment data as well as qualitative feedback in both short and long term context from the participants. As seen in Figure 1, assessment included daily feedback, multiple pre and post conceptual assessment tests for participants and their students, pre- and post-surveys of classroom practices, video as well as direct observation and evaluation of participant teaching, self-reflection journals, and finally participant evaluations of the Saturday sessions, summer institutes, and overall program.

All assessments, evaluations and surveys were identical at each administration to allow for comparison over time with the exception of a small number of questions on the final overall evaluation to address not only the past year but the entire length of program.

## Participants

A total of 35 high school science teachers participated in the 2009-2012 IPTIR program. Participant information was collected primarily via the IPTIR Participant Application collected from each participant prior to the summer session each year. The application was the same for each year of the program, and provided information such as:

- basic contact information;
- basic school information;
- number of years teaching as well as number of years teaching physics
- approximate number of students per class and number of students in school
- community population;
- status of the acquisition of their Iowa physics teaching endorsement.

Participating teachers had a range of physics teaching experience, with some starting the program with over a decade of physics teaching experience, while other participants had no physics endorsement or experience teaching physics when starting the program. Participants were from throughout the state of Iowa, representing all 9 of Iowa’s Area Education Agencies, 30 public and 3 private schools.

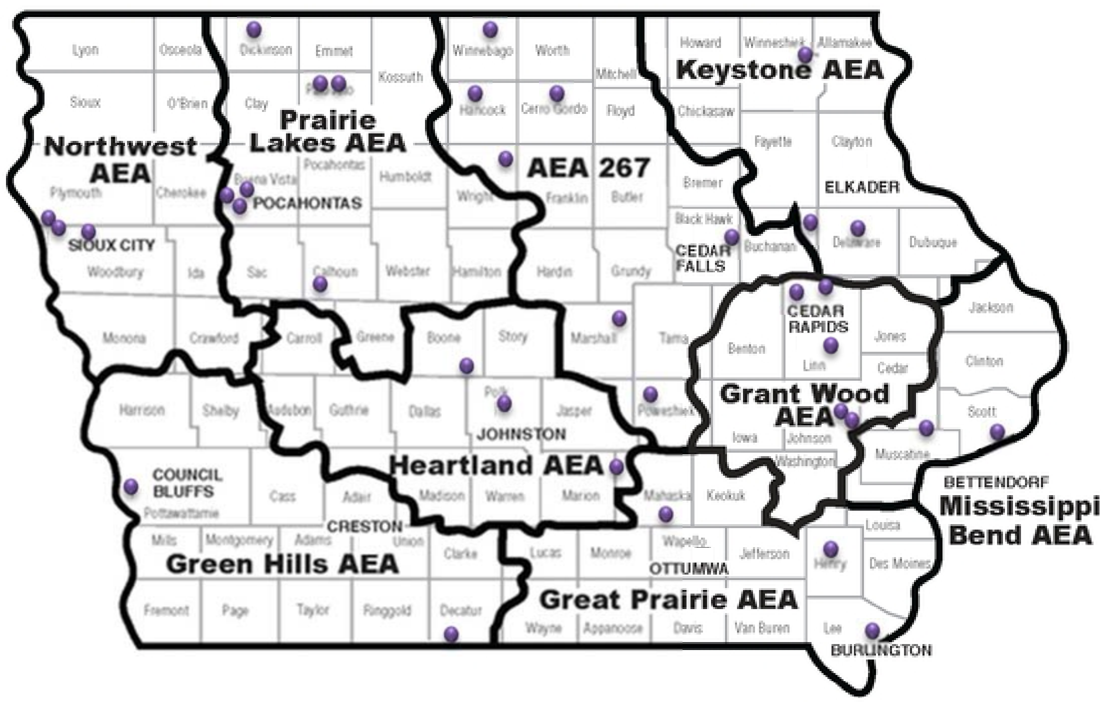



*Figure 2:* Location of IPTIR participant schools.

A breakdown of the participants by years is as follows:

<u>Year 1 participants:</u> Twenty-four participants were selected from a pool of 48 applicants based on a number of criteria including physics teaching assignments, progress towards an Iowa physics teaching endorsement, availability of classroom resources,

public vs. private school, location in the state, and district high-need status. Selected participants represented 22 public schools, 2 private schools, and 6 high needs schools.

Year 2 participants: Twenty-three of the 24 Year 1 participants reapplied and were selected for Year 2 of the IPTIR program. Additional funding made it possible to select 7 new participants to Year 2 of the program, resulting in 31 participants from all over the state of Iowa with all 9 Area Education Agencies (AEAs) represented.

Year 3 participants: Twenty-eight teachers from rural and metropolitan communities all over the state of Iowa participated in the final year of the project. Additional Title II funds were obtained to keep the teachers added in Year 2 in Year 3. Unfortunately, 6 teachers could not participate in Year 3 since they no longer met the requirements of the IPTIR program in that they did not have a physics teaching assignment. The Year 3 teachers represented 27 public schools, 4 private schools, and 5 high needs schools. All 9 AEAs were represented in the third year of the program. The Year 3 participants were similar to those of Year 1 and 2 in regard to the subjects and grade levels they were teaching as well as teaching background.

Over the course of the three-year program, 20 high-need schools were represented, defined as having at least 30% of K-12 students eligible for free and reduced lunch based on total Title V enrollment (Iowa DoE, 2012a). Refer to Appendix A for additional participant and school information, though to ensure participant privacy no directly identifying information is provided.

An overview of the participant profile is shown in Tables 1 and 2, Table 1 outlining the total number of participants in the program each year and Table 2 outlining participation based on number of years participated in the program.

Table 1

*IPTIR Participant Profile by Year*

| | Number of new participants | Number of returning participants | Year Total |
|---|---|---|---|
| Year 1<br>Summer 2009 to May 2010 | 24 | NA | 24 |
| Year 2<br>Summer 2010 to May 2011 | 8 | 23 | 31 |
| Year 3<br>Summer 2011 to May 2012 | 3 | 25 | 28 |

Table 2

*IPTIR Participant Profile by Number of Years of Participation*

| | | | | |
|---|---|---|---|---|
| Single Year Participants | 2009-10: 1 | 2010-11: 1 | 2011-12: 3 | Total: 5 |
| Two Year Participants | 2009-10 & 2010-11: 5 | 2010-11 & 2011-12: 7 | | Total: 12 |
| Three Year Participants | 2009-2010, 2010-2011, & 2011-2012 | | | Total: 18 |
| | | | | Total Cohort: 35 |

<u>Evaluation Activities</u>

Evaluation of the program utilized both formative and summative evaluation data and a combination of quantitative and qualitative methods. IPTIR faculty used formative assessment to continually adjust instruction throughout all three years of the program, and summative assessment to determine successful application of interactive engagement instructional methods, as evident by normalized gains on conceptual assessment tests and self-reported increases in content knowledge.

Conceptual Assessments

All participants were administered three conceptual assessment pretests during the first days of the summer 2009 session and again as post-tests at the end of the session. These tests were the Force Concept Inventory (Hestenes et al., 1992), the Test of Understanding Graphs - Kinematics (Beichner, 1994) and the Classroom Test of Scientific Reasoning (Lawson, 1978). The program used conceptual assessments as a diagnostic tool for teachers and students, as well as a test of evaluating effectiveness of instruction. Participating teachers were required to administer the same three conceptual and scientific reasoning tests listed above to their students. These tests were given as pre-tests and post-tests to students in their physics and physical science courses during each academic year for the duration of their involvement in IPTIR. New participants who were added to the program in Years 2 and 3 were administered the three conceptual assessments on the first days of their participating year.

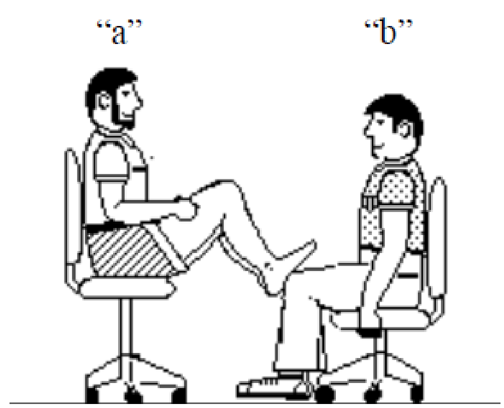
Force Concept Inventory. The Force Concept Inventory (FCI) is based on the belief that without comprehension of basic Newtonian concepts, any further instruction on force and motion will not be comprehended. The test authors (Hestenes et al., 1992) structured the questions so as to present a choice between Newtonian concepts and common alternative conceptions. The test is widely used by physics instructors to identify student misconceptions, and results are strikingly comparable among similar physics courses in both secondary and post-secondary classes (Huffman & Heller, 1995). Because reasoning is a key component of the test, something which "teaching to the test" will not necessarily improve, the FCI is valuable as a measure of teaching effectiveness (Hestenes et al., 1992).

The structure of the test is 30 multiple-choice questions covering the six major Newtonian concepts: (1) Kinematics, (2) Newton's First, (3) Second, and (4) Third Laws, (5) Superposition and (6) Kinds of Force (Hestenes et al., 1992). The questions are conceptual in nature, targeting the core ideas behind Newtonian concepts, and do not assess the use of equations or calculations. A sample problem is provided in Figure 3.

28. In the figure at right, student "a" has a mass of 95 kg and student "b" has a mass of 77 kg. They sit in identical office chairs facing each other.

Student "a" places his bare feet on the knees of student "b", as shown. Student "a" then suddenly pushes outward with his feet, causing both chairs to move.

During the push and while the students are still touching one another:




(A) neither student exerts a force on the other.
(B) student "a" exerts a force on student "b", but "b" does not exert any force on "a".
(C) each student exerts a force on the other, but "b" exerts the larger force.
(D) each student exerts a force on the other, but "a" exerts the larger force.
(E) each student exerts the same amount of force on the other.

*Figure 3.* Sample question from the Force Concept Inventory.

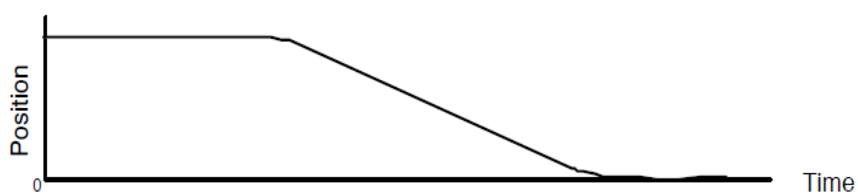
The FCI design minimizes the common problem with multiple-choice tests of false positives, where the correct, Newtonian response was chosen for incorrect reasons (Hestenes & Halloun, 1995). Incorrect responses on the FCI are structured to present common alternative conceptions that are often falsely grounded in everyday experiences. Hestenes et al. (1992) provided a taxonomy of common misconceptions associated with incorrect answers on the assessment, which give insight into the overall errors in thinking a student may possess.

The test was initially validated by the authors with over 1500 high school and over 500 university students at its inception in 1992. Follow-up interviews were also performed with students to support the test's validity. Today the test is considered one of the most consistently reliable tools for physics instructors to test for understanding of mechanics concepts (Huffman & Heller, 1995).

Hake (1998) performed a survey on interactive-engagement versus traditional teaching methods, utilizing FCI pre- and post- test results from 62 introductory physics courses (n=6542), 14 of which were identified as implementing traditional teaching methods, and 48 implementing interactive engagement methods. Hake established a normalized gain (g) in pre- to post-test score as "high-g" as those with $g \geq 0.70$, "medium-g" as $0.7 > g \geq 0.3$, and "low-g" as $g < 0.3$. This important study not only provided evidence of the effectiveness of interactive engagement methods, but also established standard measures of FCI gains that are indicative of effective teaching methods (Coletta & Phillips, 2005).

Test of Understanding Graphs in Kinematics. The Test of Understanding Graphs in Kinematics (TUG-K) is a 21-question multiple-choice test developed to uncover and address difficulties students often have interpreting kinematic graphs to understand the relationships between position, velocity, and acceleration. Graphs are an important tool in science, and used often in physics to summarize and draw conclusions from what can be a large amount of data (Beichner, 1994). Meltzer's study on the role of multiple representations in physics (2005) outlines the well-researched issue that learning difficulties and inability to master physics concepts is often due to difficulties with graphical representations. Similar to the FCI, the test is written so a choice is forced between the correct analysis and a common misinterpretation. A sample problem is provided in Figure 4.

8. Here is a graph of an object's motion. Which sentence is a correct interpretation?




(A) The object rolls along a flat surface. Then it rolls forward down a hill, and then finally stops.
(B) The object doesn't move at first. Then it rolls forward down a hill and finally stops.
(C) The object is moving at constant velocity. Then it slows down and stops.
(D) The object doesn't move at first. Then it moves backwards and then finally stops
(E) The object moves along a flat area, moves backwards down a hill, and then it keeps moving.

*Figure 4.* Sample question from the Test of Understanding Graphs – Kinematics.

The TUG-K has established validity and statistical reliability via various means, including high point-biserial coefficients, which measures how often questions answered correctly are indicative of understanding. A point-biserial coefficient greater than 0.20 is desired, and the TUG-K rated at an average coefficient of 0.74 (Beichner, 1994).

Classroom Test of Scientific Reasoning. As the name suggests, the Classroom Test of Scientific Reasoning (Lawson Test) assesses scientific reasoning skills (Lawson, 1978). The test addresses multiple dimensions in the realm of science, such as conservation of matter/volume, control of variables, deductive and inductive reasoning, proportionality and probability reasoning. The Lawson Test has 24 multiple-choice questions that are paired in terms of reasoning skills but can be scored independently. Because inquiry based teaching fosters the process of scientific thinking, this test can give insight into effective teaching independent of the specific topic being instructed. A sample question pair is provided in Figure 5.

1. Suppose you are given two clay balls of equal size and shape. The two clay balls also weigh the same. One ball is flattened into a pancake-shaped piece. *Which of these statements is correct?*

   a. The pancake-shaped piece weighs more than the ball
   b. The two pieces still weigh the same
   c. The ball weighs more than the pancake-shaped piece

2. *because*
   a. the flattened piece covers a larger area.
   b. the ball pushes down more on one spot.
   c. when something is flattened it loses weight.
   d. clay has not been added or taken away.
   e. when something is flattened it gains weight.

*Figure 5.* Sample questions from Classroom Test of Scientific Reasoning.

Pedagogical Assessment

Teacher Survey of Classroom Practice. To determine if the participants experienced a change in their teaching pedagogy, the Teacher Survey of Classroom Practices (Escalada & Stone, 2009) was administered to participants on the first day of the program and at the end of each academic year (May 2010, May 2011 and May 2012). The 91-item, self-reported survey classifies teaching and learning behaviors in terms of characteristics of teacher-centered, traditional classrooms and those which are characteristics of student-centered, interactive classrooms. Thirty-seven items on the test are associated with a more student-centered classroom, and analysis of changes in participants' scores throughout the program may be indicative of changes in their pedagogical practices.

Current research on best practice has set a goal to foster increasingly student-centered and activity based classrooms, which is consistent with IPTIR modeling of instruction and program goals. Evidence of this is a decrease in the amount of direct teaching in which the teacher presents and reviews concepts through traditional "lecturing" and where student learning is strongly tied to use of the textbook. Characteristics of student-centered classrooms in which teaching for understanding is emphasized include student reflection and discussion, engaging laboratory experiences requiring problem solving, students as decision-makers, and authentic learning experiences.

Reformed Teaching Observation Protocol. The current movement towards inquiry-based education created a need to define the aspects of reformed teaching, as well as a method to quantify the progression of this reform. Reformed teaching is rooted in the constructivist theory of learning, and therefore draws upon previous knowledge and uses interactive engagement methods to develop conceptual understanding. To define and assess these aspects of reformed teaching, the Reformed Teaching Observation Protocol (RTOP; Sawada et al., 2000) was developed to quantitatively analyze a classroom's reformed activities (Piburn & Sawada, n.d.).

Participating teachers were asked to videotape one lesson showcasing their classroom instruction during each semester of participation in the IPTIR program. In years 2 and 3 the participants were provided with video cameras that allowed the video captured to be reviewed from a more standardized format. These videos were then reviewed by IPTIR faculty and staff utilizing the RTOP assessment tool.

The 100-point RTOP is an observational instrument is broken down into five 20-point categories: Lesson Design and Implementation, Propositional Knowledge (Content), Procedural Knowledge (Inquiry), Communicative Interactions, and Student/Teacher Relationships. Each category contains five questions scored on a five-point Likert scale, with a total possible score ranging from 0 to 100 points. Higher scores reflect a greater degree of reform that is consistent with interactive engagement or student-centered techniques. Any RTOP score greater than 50 indicates considerable presence of "reformed teaching" in a lesson (MacIsaac & Falconer, 2002). The instrument is used to assess a videotaped recording of a typical classroom lesson of the teacher being evaluated. Evaluators are provided a training guide which provides clarification for each RTOP item that a score is assigned on a scale from 0 to 4 (Sawada

& Piburn, 2000). The IPTIR faculty and staff used an established RTOP scoring form to score each video. An example from the RTOP scoring form is shown in Figure 6.

| III. | LESSON DESIGN AND IMPLEMENTATION | Never Occurred | | | | Very Descriptive |
|---|---|---|---|---|---|---|
| 1) | The instructional strategies and activities respected students' prior knowledge and the preconceptions inherent therein. | 0 | 1 | 2 | 3 | 4 |
| 2) | The lesson was designed to engage students as members of a learning community. | 0 | 1 | 2 | 3 | 4 |
| 3) | In this lesson, student exploration preceded formal presentation. | 0 | 1 | 2 | 3 | 4 |
| 4) | This lesson encouraged students to seek and value alternative modes of investigation or of problem solving. | 0 | 1 | 2 | 3 | 4 |
| 5) | The focus and direction of the lesson was often determined by ideas originating with students. | 0 | 1 | 2 | 3 | 4 |

*Figure 6.* Sample from the Reformed Teaching Observation Protocol scoring form.

The nature of the tool does present limitations that should be taken into account. While useful for assessing elements of the videotaped lesson, each scoring event deals with only one class, which may or may not be indicative of a "typical" class period. The single class aspect also leads to inconsistencies comparing scores among a population of teachers or change over time of an individual teacher. Unless the same lesson is being executed in all videos, it is difficult to isolate differences between the teachers versus differences within the nature of the lessons. Additionally, variability within the perceptions of the scorers can be problematic due to the test's interpretive nature.

Inter-rater reliability was established by Sawada and Piburn (2000) using an evaluating team which reviewed tapes using the RTOP. The judgments of the reviewers were discussed, and the interpretations developed during this process were compiled into the RTOP training guide to increase evaluator reliability. A sample from the RTOP training guide is provided in Figure 7. The RTOP has shown to be a valuable tool in assessing inquiry based pedagogical skills and has been used in previous studies to provide empirical data showing a correlation between content knowledge and the degree a teacher utilizes the reform method of teaching (Park, Jang, Chen & Jung, 2010).

III. LESSON DESIGN AND IMPLEMENTATION

**1) The instructional strategies and activities respected students' prior knowledge and the preconceptions inherent therein.**

A cornerstone of reformed teaching is taking into consideration the prior knowledge that students bring with them. The term "respected" is pivotal in this item. It suggests an attitude of curiosity on the teacher's part, an active solicitation of student ideas, and an understanding that much of what a student brings to the mathematics or science classroom is strongly shaped and conditioned by their everyday experiences.

**comments:**

**2) The lesson was designed to engage students as members of a learning community.**

Much knowledge is socially constructed. The setting within which this occurs has been called a "learning community." The use of the term community in the phrase "the scientific community" (a "self-governing" body) is similar to the way it is intended in this item. Students participate actively, their participation is integral to the actions of the community, and knowledge is negotiated within the community. It is important to remember that a group of learners does not necessarily constitute a "learning community."

**comments:**

*Figure 7.* Sample from the Reformed Teaching Observation Protocol training guide.

All IPTIR faculty and staff underwent RTOP training to ensure validity in scoring the videotapes. Faculty and staff individually watched and assessed two different videos of a university faculty member teaching a physics class. After watching and scoring each video, the faculty and staff would meet and share their evaluations, comparing them with completed scoring forms for each video provided with the RTOP training guide. The training videos included both traditional and reformed instruction to provide practice evaluating teaching from both ends of the pedagogical spectrum. The evaluator training was executed with the intent of arriving at a consensus on the interpretations of specific teaching indicators and how to score consistently.

Program Evaluation

To determine if the program was meeting the needs of the participants, IPITR faculty gathered feedback from the participants using a program evaluation survey daily during the summer sessions, after each Saturday session, and the end of each academic year.

During the summer sessions, daily feedback was used to continually adjust instruction throughout all three summers of the IPTIR program. The daily feedback evaluation tool asked participants to indicate whether each portion of the day was helpful, if participant needed more time to understand the ideas presented, needed more practice, needed more examples or if already know most of what was covered. The form also included a comment portion for the participants to elaborate on their indicated ratings. These forms were completed at the end of each day, and the previous day's feedback was reviewed on a daily basis by IPTIR faculty at morning meetings. Instruction was then adjusted to address concerns and questions raised by participants on their daily feedback forms.

At the conclusion of the summer session program participants evaluated their overall summer workshop experience. In addition to providing insight into structuring the subsequent summer workshops, in summative terms this data provides evidence of the success of IPTIR in addressing the needs of the IPTIR participants.

Participant feedback surveys were collected at the end of each academic year Saturday session during all three years of the project.  The tool had the same format as the Daily Feedback form, and participant responses were used in planning upcoming sessions and subsequent summer session.

In May of each academic year participants were asked to complete an IPTIR Project Evaluation, designed by the program developers, which provided participant feedback on the program as a whole. This evaluation included questions addressing the teachers' perception of the effectiveness of the overall design and activities of the IPTIR program.

# CHAPTER 4

# RESULTS

To determine the degree to which the structure and methods of the Iowa Physics Teacher Instruction and Resources (IPTIR) program attained its principle objective of improving the teaching quality of its participants, data was collected throughout the duration of the IPTIR program from multiple sources. Changes in participant content knowledge were analyzed using conceptual pre- and post-test scores, as well as self-reported changes in physics knowledge. Changes in participant pedagogy were analyzed using self-reported Teaching Survey of Classroom Practices, and these results are supported by the faculty-assessed Reformed Teaching Observation Protocol (RTOP). Conceptual assessments completed by students in the participants' classrooms were analyzed as evidence of an improvement in the participants' quality of physics teaching.

In order to isolate the impact of the full three-year program, only data from participants that participated in all three years of the program (n=17) was used in analysis, unless indicated otherwise. Normalized gain was calculated from averages of the pre- and post-test scores for the conceptual assessments of both the three-year participants and their students. Normalized gain is considered a meaningful measure of how well the program increased participant content knowledge (Coletta & Phillips, 2005). The Normalized gain, $G$, is a measure of the change in score from pretest to posttest, divided by the maximum possible increase based on the pretest score.

$$Average\ Normalized\ Gain\ (G) = \frac{(\%\ post - \%\ pre)}{(100 - \%\ pre)}$$

Changes in teaching pedagogy were based on calculated change of the Likert scale responses from the pre- to post-survey results, and average changes in RTOP scores.

Correlations were investigated to further isolate any contributing factors which may have affected teacher content knowledge, teacher pedagogy, and student content knowledge, as well as any correlations among these three indicators. Correlation was determined using a Pearson product-moment correlation coefficient, with significance determined using a two-tailed t-test.

## Changes in Participant Content Knowledge

### Participant Conceptual Assessments

All participants were administered three conceptual assessments, the Force Concept Inventory (FCI), Test of Understanding Graphs- Kinetics (TUG-K), and Classroom Test of Scientific Reasoning (Lawson Test), as pre-tests during the first days of the summer 2009 session and again at the end of the summer 2009 session. Pre- and post-test scores of those who participated in the entire program were averaged and the average gain for each conceptual assessment was calculated (Table 3).

Table 3

*Three-year Teacher Participant Conceptual Assessment Average Scores and Gains*

| | Pre-test | Post-test | Average Gain | STD DEV |
|---|---|---|---|---|
| FCI | 56.9% | 69.8% | 0.30 | 0.22 |
| TUG-K | 51.5% | 73.7% | 0.46 | 0.21 |
| Lawson Test | 81.9% | 81.4% | 0.01* | 0.40 |

*Note:* *Because the Lawson Test scores decreased from pre- to post-test, normalized *change* was calculated from maximum possible *decrease* (Marx & Cummings, 2007).

Self-Reported Content Knowledge

The participant's self-reported outcomes of changes in content knowledge were provided via the End of Year Program Evaluations. At the conclusion of the program, the entire cohort of participants indicated they "strongly agree" or "agree" that their participation in the IPTIR program resulted in increased awareness or updating of knowledge of physics/physical science concepts and pedagogy related to the courses taught or plan to teach. Table 4 lists the average self-reported scores from the entire IPTIR cohort of questions pertaining specifically to conceptual gains. While the increases are not statistically significant, the responses demonstrate the participants consistently perceived gains in content knowledge and conceptual understanding/learning throughout the three years of the program.

Table 4

*Participant Evaluation Scores Pertaining to Gains in Conceptual Knowledge*

| | Year 1 (n=19) | Year 2 (n=31) | Year 3 (n=22) | Avg STD DEV |
|---|---|---|---|---|
| The overall design and activities:<br>16. Provided opportunities that result in gaining knowledge about physics content and pedagogy. | 4.60 | 4.63 | 4.79 | 0.46 |
| My participation in this program has led to:<br>43. Increased awareness or updating of my knowledge of physics/physical science concepts and pedagogy related to the courses I teach or plan to teach. | 4.74 | 4.77 | 4.79 | 0.45 |

*Note*: Likert scale responses ranged from 1 (strongly disagree) to 5 (strongly agree).
Avg STD DEV = Average of the Standard Deviations of responses for each year.

Changes in Participant Pedagogy

The Teacher Survey of Classroom Practices (TSCP) provided evidence of changes in participants' teaching pedagogy. This instrument was administered as a pre-test on the first day of the IPTIR summer institute in Year 1 (2009) and again in May of 2010, May of 2011, and in May of 2012. Analysis of this 91-item survey classified teaching and learning behaviors in terms of teaching strategies characteristic of traditional, teacher centered classrooms and those which are characteristic of interactive, student-centered classrooms. Current research on best practice sets the goal of increasingly student- centered and activity based classrooms, which is consistent with IPTIR modeling of instruction and program goals.

Evidence of shifts towards research-based methods is a decrease in the amount of direct teaching in which the teacher presents and reviews concepts through traditional "lecturing" and student learning is strongly tied to use of the textbook. A decrease in teacher-centered characteristics is ideally paired with an increase in characteristics of student-centered classrooms in which teaching for understanding is emphasized, including student reflection and discussion, problem solving, students as decision-makers, and interactive engagement methods.

Select items were identified on the TSCP as strong indicators of a teacher-centered classroom (11 items) and others as indicators of a more student-centered classroom (26 items). Individual changes in pre- to post-survey responses were calculated, and average cohort changes were calculated based on these individual changes. Participants' average responses to the teacher-centered items decreased from pre- to post-test assessments, and average responses to the student centered items increased from pre- to post-test. A sample of the total average participant change in the frequency of teacher-centered and student-centered characteristics is provided in Tables 5 and 6.

Table 5

*Select Items Representing Characteristic of Teacher-Centered Classrooms*

| | Pre-Test (n=17) | Year 1 Post (n=16) | Year 2 Post (n=17) | Year 3 Post (n=14) | **Avg Total Change** | Avg STD DEV |
|---|---|---|---|---|---|---|
| 22. Answer questions and/or solve problems from a textbook or worksheet. | 4.35 | 3.44 | 3.59 | 3.69 | **-0.66** | 0.74 |
| 38. Listen to the teacher explain something about physics. | 3.18 | 2.00 | 2.18 | 2.54 | **-0.64** | 0.76 |
| 39. Read about physics from a textbook | 2.29 | 1.31 | 1.59 | 1.77 | **-0.52** | 0.89 |
| 46. Students followed step-by-step instructions. | 1.88 | 1.44 | 1.35 | 1.38 | **-0.50** | 0.46 |

*Note*: Likert scale responses range from 1 (Never) to 5 (3-5 times per week).
Avg STD DEV = Average of the Standard Deviations of responses for each year.

Table 6

*Select Items Representing Characteristic of Student-Centered Classrooms*

| | Pre-Test (n=17) | Year 1 Post (n=16) | Year 2 Post (n=17) | Year 3 Post (n=14) | **Avg Total Change** | Avg STD DEV |
|---|---|---|---|---|---|---|
| 16. Students reflect on their own learning. | 2.18 | 3.07 | 3.00 | 3.57 | **+1.39** | 1.23 |
| 18. Work in pairs or small groups. | 4.35 | 4.50 | 4.65 | 4.79 | **+0.44** | 0.68 |
| 27. Take part in group or class discussion. | 3.82 | 4.44 | 4.24 | 4.36 | **+0.52** | 0.81 |
| 28. Change something in an experiment to see its effects. | 2.71 | 3.38 | 3.53 | 3.36 | **+0.65** | 0.82 |
| 29. Design experiments. | 2.00 | 3.19 | 3.06 | 3.00 | **+1.00** | 0.73 |
| 34. Choose a method for expressing an idea to the class. | 2.12 | 2.94 | 2.76 | 3.00 | **+0.88** | 1.13 |
| 35. Revise and improve student's own work. | 2.18 | 3.06 | 3.18 | 3.21 | **+1.03** | 1.13 |

*Note*: Likert scale responses range from 1 (Never) to 5 (3-5 times per week).
Avg STD DEV = Average of the Standard Deviations of responses for each year.

A broad analysis of changes in participant teaching pedagogy over the three years of the program is shown in Table 7. Responses of participants of all three years were averaged for all student -centered measures (26 items) as well as all teacher centered measures (11 items). Participants continued to change in a positive direction with each additional year, overall decreasing in teacher-centered practices and increasing in student-centered.

Table 7

*Average Frequencies and Changes in Participant Teaching Pedagogy*

| | Pre-Test (n=17) | Year 1 Post (n=16) | Year 2 Post (n=17) | Year 3 Post (n=14) | **Avg Total Change** | STD DEV |
|---|---|---|---|---|---|---|
| Student centered characteristics | 2.74 | 3.30 | 3.29 | 3.26 | **+0.55** | 0.31 |
| Teacher centered characteristics | 2.81 | 2.16 | 2.37 | 2.40 | **-0.37** | 0.45 |

*Note*: Likert scale responses range from 1 (Never) to 5 (3-5 times per week).
STD DEV = Standard Deviation of the Average Total Change for each participant.

The self-reported changes in teaching pedagogy are supported by videotape analysis of one video lesson showcasing participants' classroom instruction. Participants were asked to provide a video each semester of participation in the IPTIR program, but not all participants submitted videos, resulting in a smaller sample size for analysis. The videotapes were viewed and scored by the IPTIR faculty and staff using the modified Reformed Teaching Observation Protocol (RTOP; appended).

Average scores of three-year participants who submitted videos started below the "reformed teaching" indicator of 50 points (MacIsaac & Falconer, 2002), and progressed steadily throughout the three years of the program, supporting the participant's self-reported changes in pedagogy. Further analysis shows the greatest improvements were in procedural knowledge (31.2% increase) and student/ teacher interactions (27.9% increase), as show in in Table 8.

Table 8

*Average Three-Year Participant RTOP Scores*

| | Fall 2009 (n=8) | Spring 2010 (n=5) | Spring 2011 (n=12) | Spring 2012 (n=7) | **Overall % Change** | STD DEV |
|---|---|---|---|---|---|---|
| Average RTOP Total score (max 100) | 36.2 | 40.4 | 58.3 | 64.2 | **23.8 %** | 15.8 % |
| *Individual Category Scores (max 20 points per category)* | | | | | | |
| Lesson Design and Implementation | 6.6 | 7.7 | 10.7 | 12.0 | **27.2 %** | 22.5 % |
| Content: Propositional Knowledge | 8.6 | 10.4 | 12.9 | 14.0 | **18.2 %** | 28.4 % |
| Content: Procedural Knowledge | 4.7 | 4.1 | 9.0 | 10.3 | **31.2 %** | 23.7 % |
| Classroom Culture: Communicative Interactions | 8.0 | 8.5 | 11.8 | 12.5 | **20.1 %** | 19.0 % |
| Classroom Culture: Student/Teacher Relationships | 8.6 | 9.7 | 13.9 | 15.3 | **27.9 %** | 17.9 % |

Changes in Participants' Student Achievement

All IPTIR Participants were asked to administer to their physics/physical science students the same three conceptual assessments tests taken by the participants, the Force Concept Inventory (FCI), Test of Understanding Graphs- Kinetics TUG-K), and Classroom Test of Scientific Reasoning (Lawson Test). Student learning was assessed by analyzing improvement from pre- to post-test by calculating the average normalized gain, which is based on how much room for improvement lies in each student's pretest score. Table 9 summarizes the average student score for each year as well as average and total gains.

Table 9

*Average Student Conceptual Assessment Post-Test Scores and Gains*

| | Year 1 | | Year 2 | | Year 3 | |
|---|---|---|---|---|---|---|
| | Avg Score | Avg Gain | Avg Score | Avg Gain | Avg Score | Avg Gain |
| FCI | 41.5% (n=383) | 0.20 | 47.7% (n=596) | 0.40 | 48.1% (n=466) | 0.40 |
| TUG-K | 48.0% (n=387) | 0.34 | 53.9% (n=604) | 0.37 | 56.0% (n=496) | 0.40 |
| Lawson Test | 67.0% (n=389) | 0.19 | 65.8% (n=598) | 0.24 | 71.3% (n=425) | 0.22 |

## Correlations

Correlations were investigated to determine if factors such as general teaching experience and/or physics-specific teaching experience contributed to the evaluation results. Additionally, results of the evaluation were compared against each other to investigate if any program outcomes correlated to each other. A Pearson product-moment correlation coefficient was calculated to determine the relationship between the variables, and statistical significance was determined using a two-tailed t-test.

### Participant Content Knowledge

Because the FCI and TUG-K directly assess content knowledge, whereas the Lawson Test assesses scientific reasoning, investigations into content knowledge will be limited to the FCI and TUG-K results. In the cohort of teacher participants in the program all three years, there were two teachers that had been teaching physics for more than five years, with 15 and 19 years of physics teaching experience. Both of these teachers performed very well on the conceptual assessments, and with the inclusion of their data there is a statistically significant correlation between the number of years teaching physics and both teacher pre-test performance on the FCI ($r(15) = 0.684$, $p < 0.01$) and teacher post-test performance on the FCI ($r(15) = 0.660$, $p < 0.01$). Similar correlations were found between physics teaching experience and teacher pre-test performance on the TUG-K ($r(15) = 0.618$, $p < 0.01$) (Figure 8), as well as teacher post-test performance on the TUG-K ($r(15) = 0.574$, $p < 0.02$) (Figure 9).

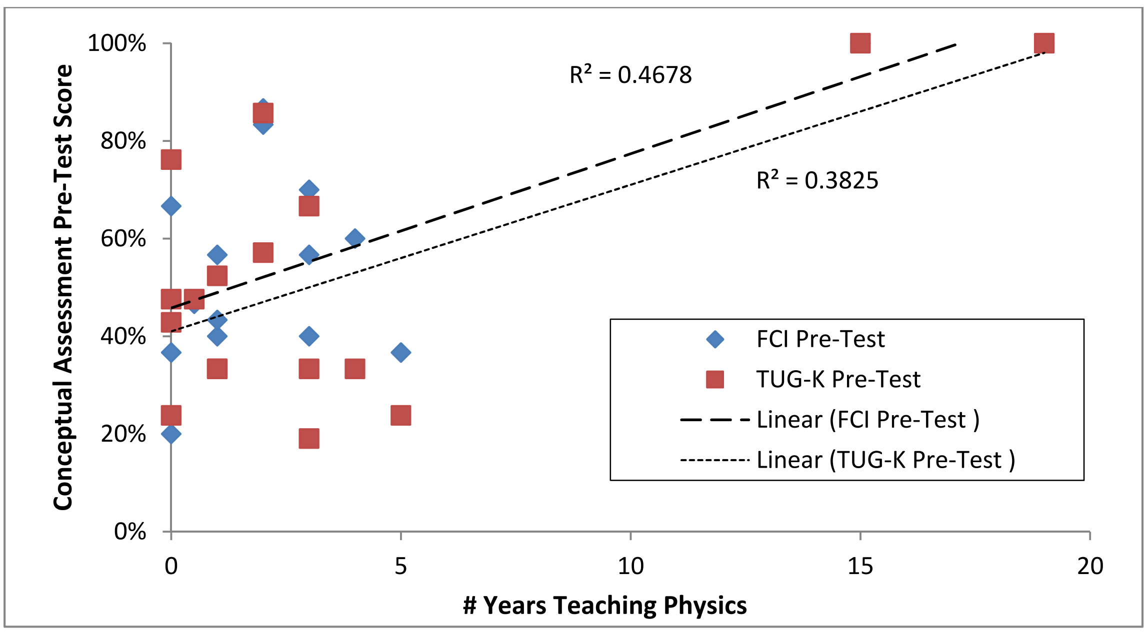



*Figure 8.* Correlation of teacher conceptual assessment pre-test scores with physics teaching experience.

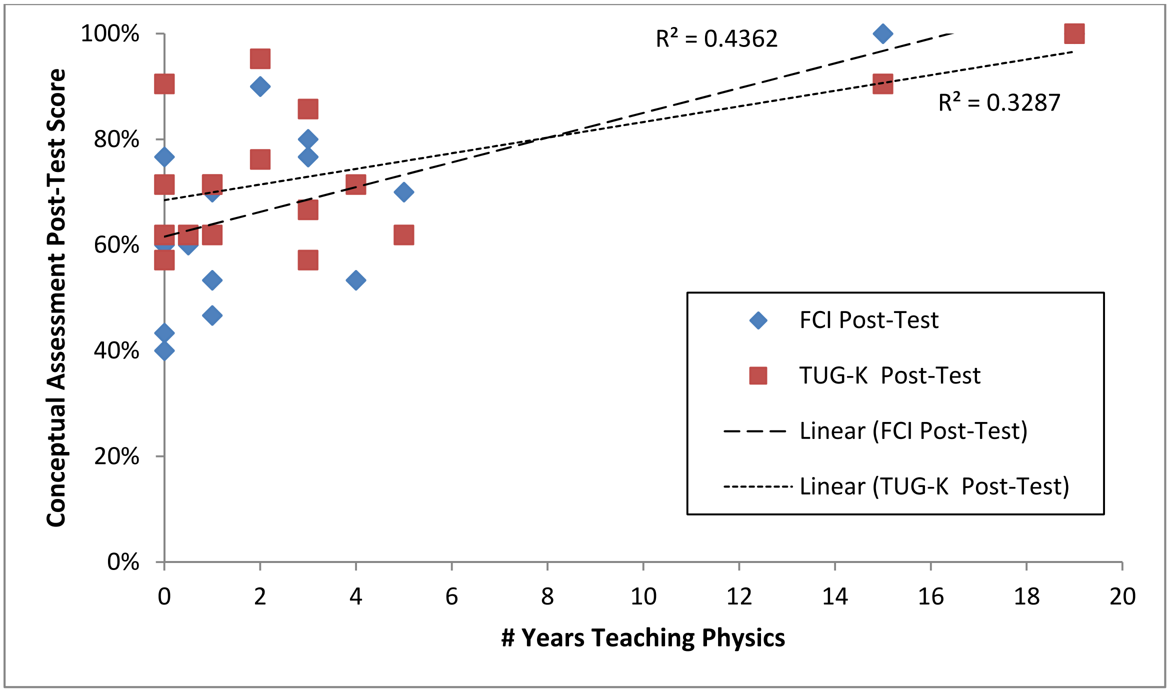



*Figure 9.* Correlation of teacher conceptual assessment post-test scores with physics teaching experience.

Because there are no participants in the full three-year cohort that had between 6 and 15 years of physics teaching experienced, it cannot be determined if the two teachers with much greater experience are outliers or if they indicate a trend that would continue as teaching experience increases from six to fifteen years. When analyzing only the participants that have five or fewer years of physics teaching experience, the correlation with conceptual assessment performance is no longer significant, as shown in Figure 10 and Figure 11.

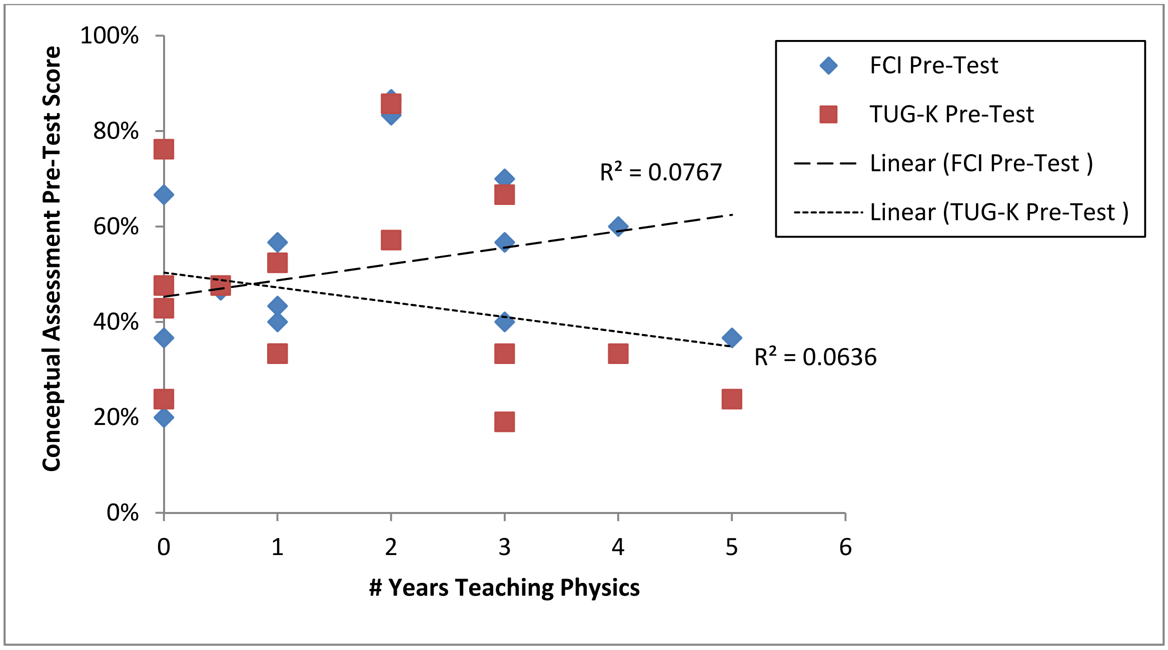



*Figure 10.* Correlation of teacher conceptual assessment pre-test scores with physics teaching experience of participants with five or fewer years of physics teaching experience.

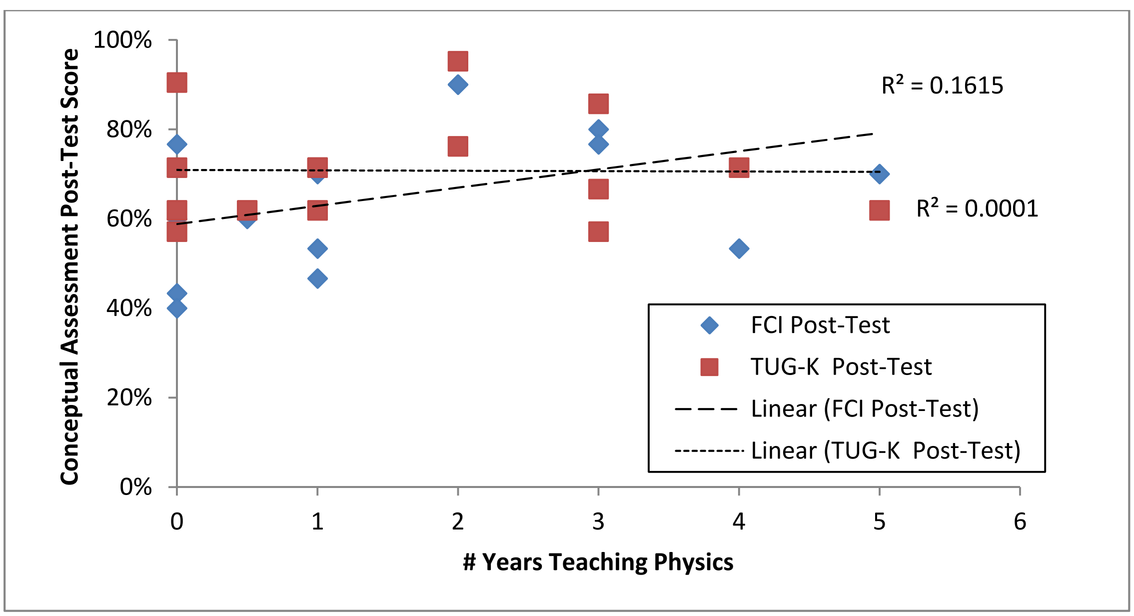



*Figure 11.* Correlation of teacher conceptual assessment post-test scores with physics teaching experience of participants with five or fewer years of physics teaching experience.

In terms of general teaching experience not specific to physics, there were no significant correlations between teaching experience and participant conceptual assessment performance. Additionally, no correlations were found between the normalized gain from pre- to post-test scores and teaching experience.

Few factors appeared to correlated with teacher' scientific reasoning skills, as indicated by the Lawson Test post-test student performance and the gain from teacher pre- to post-test. There were no significant correlations to general teaching experience. When the two participants with 15 and 19 years of physics teaching experience were included in the data set, there was a significant correlation with physics teaching experience and teacher Lawson Test post-test score ($r(15) = 0.416$, $p < 0.1$), but as with the FCI and TUG-K the two data points greatly influenced the correlation. When omitting the two participants with more than five years of physics teaching experience and analyzing only the remaining cohort, which contained only teachers with five or fewer years of physics teaching experience, there was very little to no correlation of physics teaching experience with teachers science reasoning skills ($r(13) = 0.036$). Additionally, no correlations were found between the teachers' science reasoning performance and the frequency of the teachers' student-centered characteristics or the change in frequency of the teachers' student-centered characteristics.

Correlations were found among the teachers' scores from all three conceptual assessments, as indicated between the Lawson Test and FCI post-test scores ($r(15) = 0.573$, $p < 0.02$), the Lawson Test and TUG-K post-test scores ($r(15) = 0.523$, $p < 0.05$), shown in Figure 12, and between the FCI and TUG-K ($r(15) = 0.476$, $p < 0.10$), shown in Figure 13.

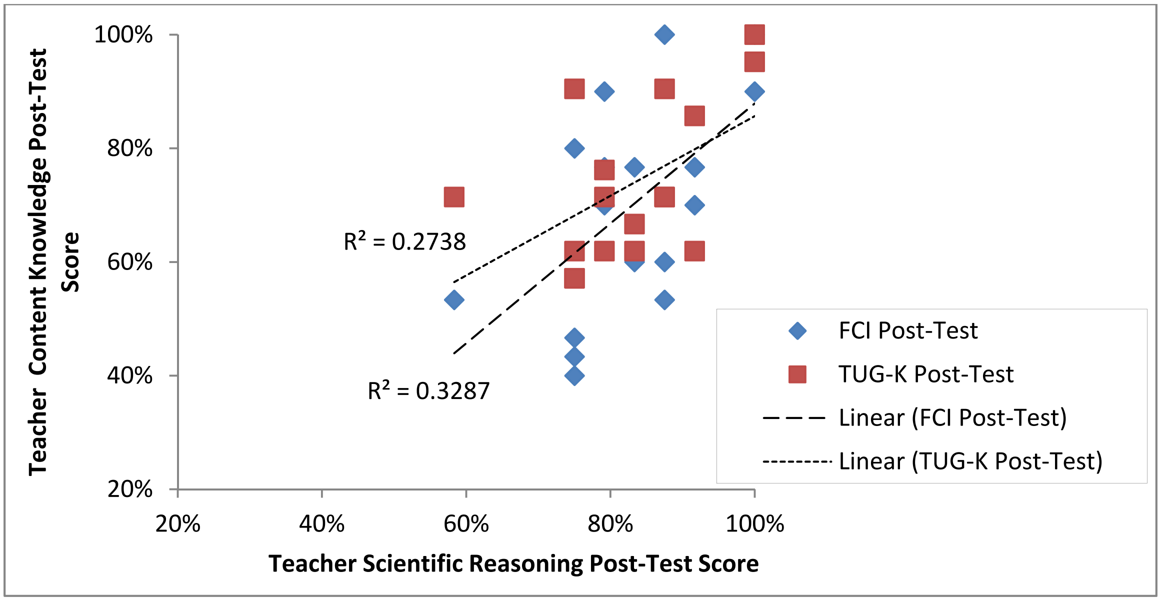



*Figure 12.* Correlation of teacher content knowledge with teacher scientific reasoning.

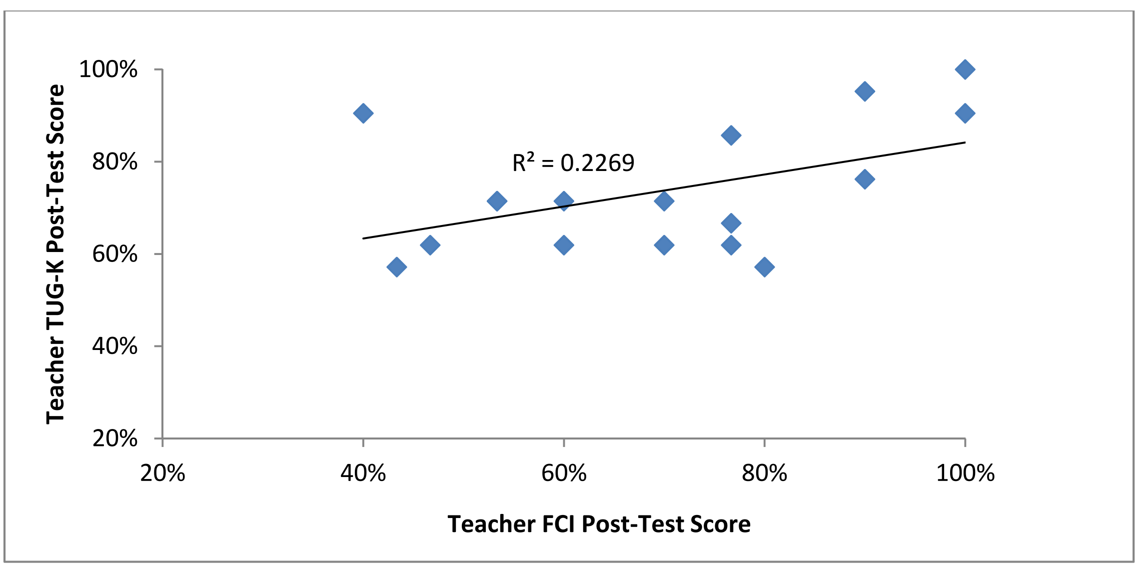



*Figure 13.* Correlation of teacher FCI post-test scores with teacher TUG-K post-test scores.

Teaching Pedagogy

Teaching experience did not appear to have an impact on the participants' initial teaching characteristics, nor overall change in the self-reported frequency of student-centered characteristics in their classrooms. There was a statistically significant negative correlation between physics-specific teaching experience and the pre-test frequency of teacher-centered characteristics (r(15) = -0.439, $p < 0.10$), as shown in Figure 14**.** Degree of frequency is based on a Likert scale response of 1 (Never) to 5 (3-5 times per week). The correlation between general teaching experience and the pre-test frequency of teacher-centered classrooms was also negative, however not statistically significant.

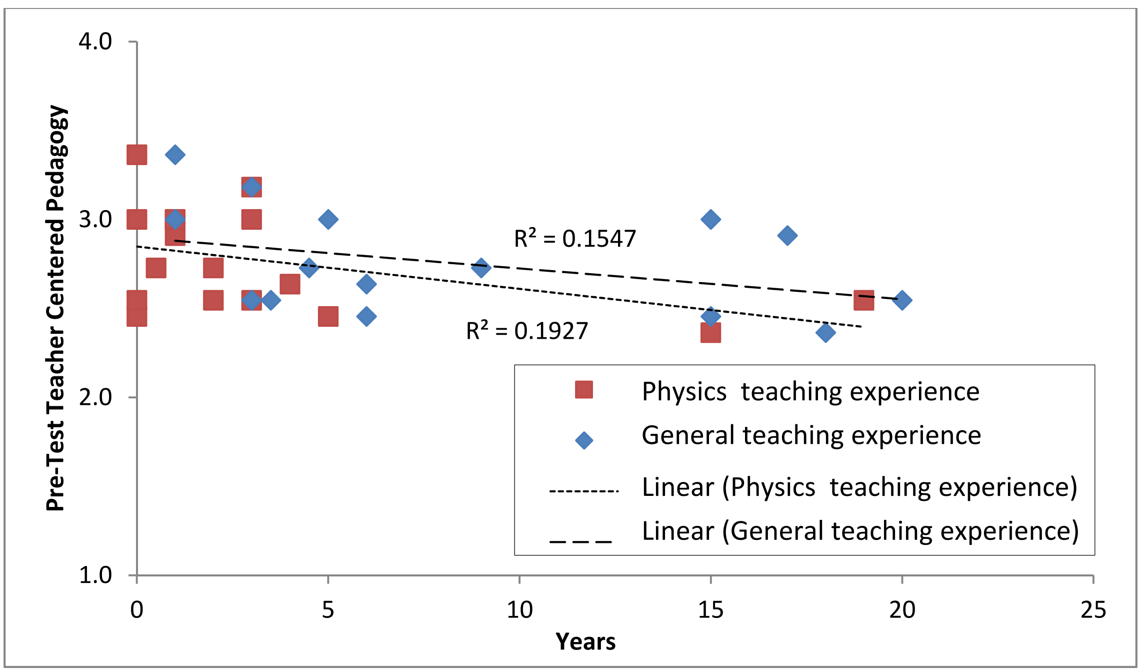



*Figure 14.* Correlation of general and content-specific teaching experience with initial self-reported frequency of teacher-centered characteristics.

There was a significant correlation between the degree of frequency of teacher-centered characteristics at the beginning of the program and the overall change in degree of frequency of student-centered characteristics ($r(17) = 0.414$, $p < 0.10$) as shown in Figure 15.

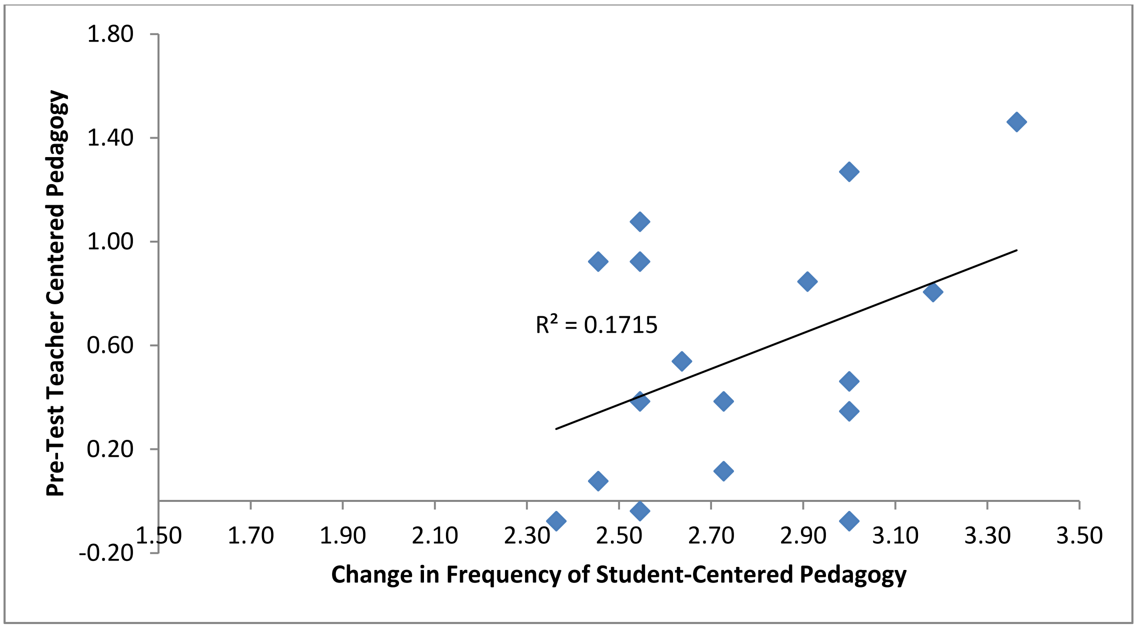



*Figure 15.* Correlation of the frequency of teacher-centered characteristics at the beginning of the program with overall change in frequency of student-centered characteristics.

No significant correlations were found between teaching experience and frequency of teacher centered practices at the end of the program, nor the total change of reported frequency. Additional there was not a significant correlation between change in frequencies of teacher-centered characteristics and student-centered characteristics.

Student Achievement

While more effective teaching is the penultimate goal of any professional development program, the ultimate goal is increased student learning. Student learning was indicated by both average post-test scores as well as average pre- to post-test gains on the conceptual assessments administered to the students at the beginning and end of each academic year.

Participants' post-test performance on the FCI conceptual assessments significantly correlated with both their student's post-test FCI scores (r(14) = 0.443, $p < 0.10$), as shown in Figure 16, and the student FCI score gains (r(14) = 0.513, $p < 0.05$), as shown in Figure 17. Similar correlations were not found with the teachers' pre-test scores.

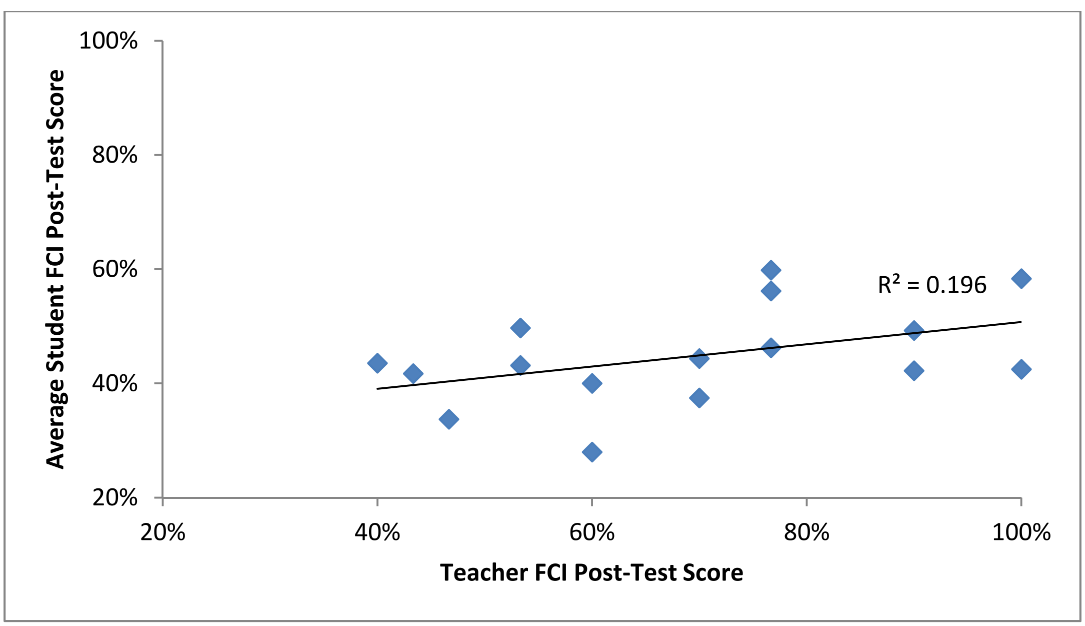



*Figure 16.* Correlation of average student FCI post-test score with FCI post-test score of teacher.

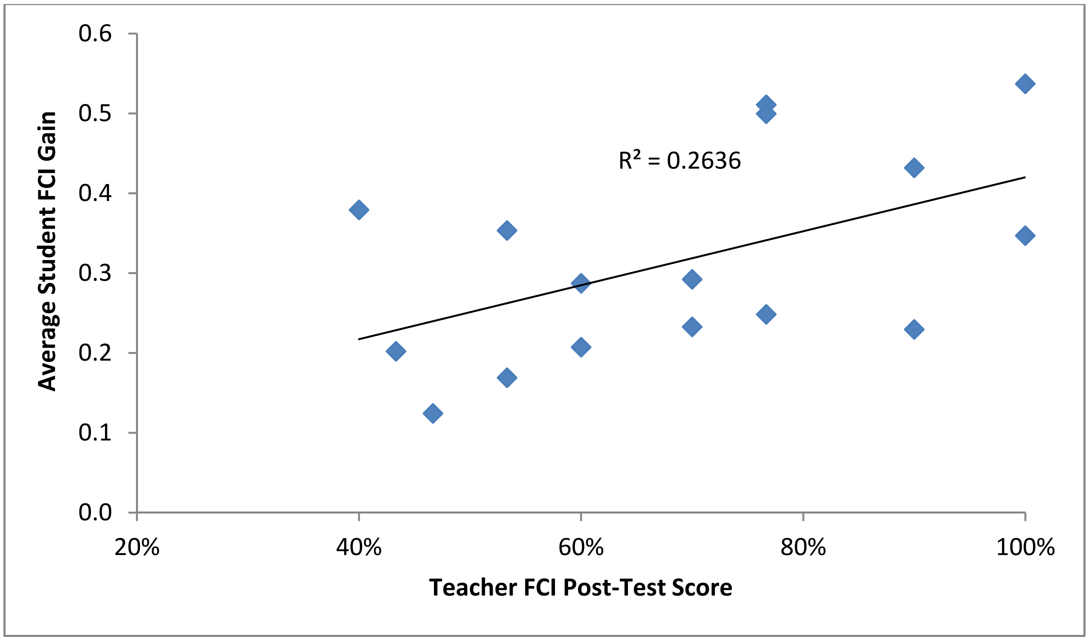


*Figure 17.* Correlation of average student FCI gain with FCI post-test score of teacher.

Content-specific experience seemed to also be a factor in student performance and learning. There was a significant correlation between physics teaching experience and the student FCI pre- to post-test gain ($r(14) = 0.492$, $p < 0.10$), as shown in Figure 18**.** This similar correlation was not found between general teaching experience and student FCI pre- to post-test gain. When analyzing the data from only teachers with five or fewer years of physics teaching experience, the correlation with student FCI gain was still present though not statistically significant ($r(12) = 0.363$).

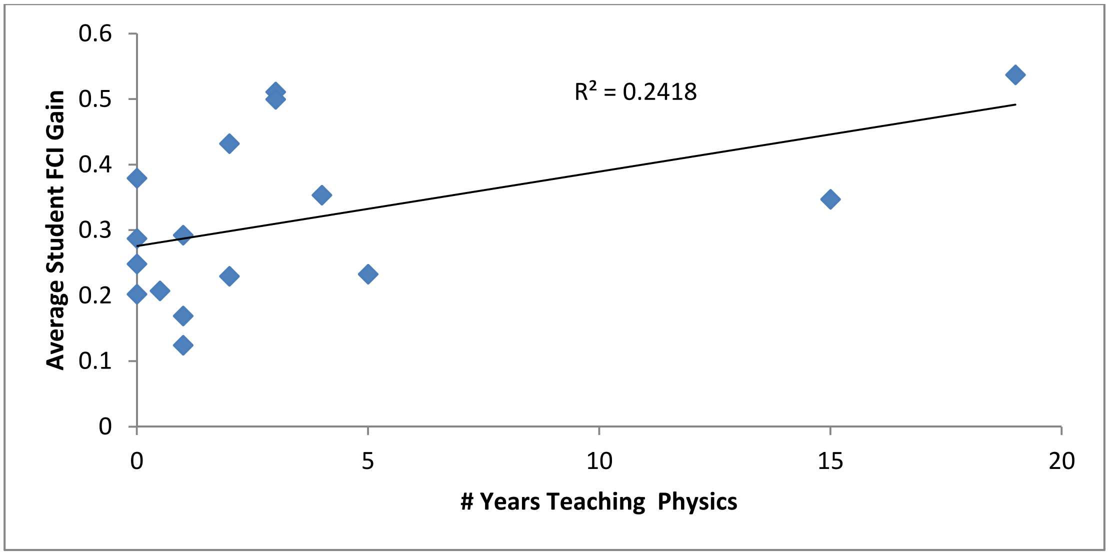



*Figure 18.* Correlation of student FCI score gain with physics teaching experience.

Significant correlation was calculated between the participants' average student TUG-K post-test scores and the average student FCI post-test scores (r(14) = 0.690, $p < 0.01$), as shown in Figure 19**.**

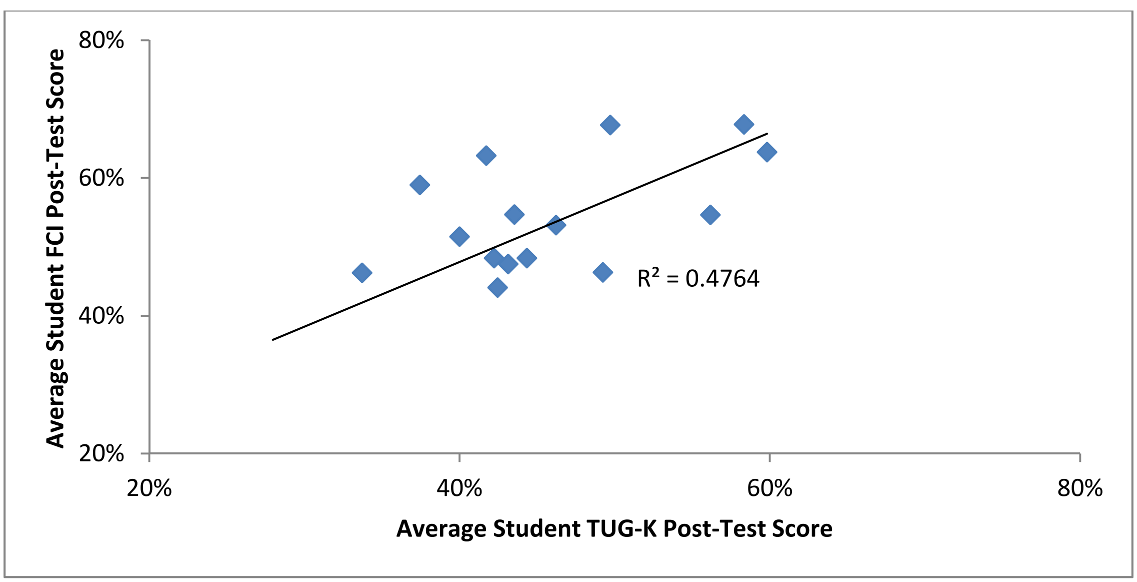



*Figure 19.* Correlation of student TUG-K Post Test score with student FCI post-test score.

Correlation was analyzed between student performance and the average high-needs status of the participant's schools, with high-needs defined as at least 30% of K-12 students eligible for free and reduced lunch based on total Title V enrollment (Iowa DoE, 2012d). No significant correlations were found between high-needs status and student scores or gains on any of the conceptual assessments. Lastly, no significant correlations were found between student performance and participants' frequency of student-centered practices, or the frequency of teacher-centered practices.

# CHAPTER 5

# CONCLUSIONS

The objective of the Iowa Physics Teacher Instructional Resources (IPTIR) program was to prepare more high-quality high school physics teachers for Iowa schools and consequently improve the classroom performance of their students. The intent of this study was to determine if this program effectively achieved this objective and to what degree, as well as to investigate what factors of the program contributed to any successful outcomes. Findings of the study provide evidence-based recommendations for much needed future professional development programs to effectively prepare high-quality high school physics teachers.

Research for this study focused on the evaluation of participant content knowledge, participant teaching pedagogy, and content knowledge of the participants' students. Evaluation of the program utilized a combination of qualitative and quantitative methods related obtaining the program's objectives. Conceptual assessments were utilized to determine if the program increased participant content knowledge and scientific reasoning skills, and conceptual assessments were also used to evaluate learning of students in the participants' classrooms. Participants were administered surveys of classroom practices to provide insight into changes in participants' teaching pedagogy, data which was supported by video analysis of classroom practices.

## Findings

With the conclusion of the three-year program, evidence indicates the program was successful developing the teacher participants' physics content knowledge and pedagogy with the focus on the latest research-based instructional methods. The participant and student conceptual assessment gains, as well as trends in participants' teaching pedagogy towards increasingly interactive classrooms demonstrate the successful implementation of the methods utilized in the IPTIR program. Conclusions provide insight to the previously posed fundamental questions of this study:

1. What changes occur throughout the professional development program in regards to participant content knowledge?
2. Did the participants modify their teaching pedagogy in alignment with today's research-based instructional methods?
3. To what degree, if any, did changes in teacher content knowledge and/or pedagogy affect the classroom performance of the participants' students?

### Changes in Teacher Participant Content Knowledge

Conceptual assessments indicated successful application of the program's interactive engagement instructional methods, evidenced by normalized gains on conceptual assessment tests and self-reported increases in participant content knowledge. The normalized gain calculated from the average scores of participants of all three years of the program for both the Force Concept Inventory (FCI) and Test for Understanding Graphs- Kinematics (TUG-K) indicates significant improvement in the teacher participants' conceptual understanding of physics. The mean pre-test score for the teachers' Classroom Test of Scientific Reasoning (Lawson Test) was much higher than those of the FCI and TUG-K pre-tests. The Lawson Test had a slight drop in average score from pre- to post-test, but the change was very minimal with an average drop in score of only 0.3% and calculated normalized change of 0.01. It can be interpreted that many science teachers have reasonably strong science reasoning skills, therefore the two weeks of instruction between pre- and post-test was not enough time to further impact these skills.

When compared to Hake's study of interactive-engagement (IE) vs traditional methods (1998), the average teacher FCI gain of 0.30 would fall under Hake's definition of "Medium gain." Hake's average gain for a traditional classroom was $0.23 \pm 0.04$, as compared to average gain of an IE classroom of $0.48 \pm 0.14$. IPTIR's utilization of IE methods and normalized gain results support Hake's conclusion that IE methods enhance problem-solving ability as compared to courses which utilize traditional teaching methods. Hake considers average gains above 0.30 to indicate successful application of interactive engagement instructional methods (Hake, 1998), which IPTIR achieved.

Hestenes and Halloun (1995) established an FCI score of 85% as the threshold for mastery of Newtonian mechanics, and 60% as the threshold for beginning Newtonian reasoning. Three of the four participants that achieved 85% mastery on the FCI post-test also scored above 85% on the pre-test, and as such in only one instance did the program effectively bring a participant from below to above Newtonian mastery. Six of the twelve participants that achieved 60% mastery scored above 60% on the pre-test as well, with six participants improving from below to above the threshold of beginning Newtonian reasoning. The teacher participants on average did not achieve the 85% threshold for mastery, but did on average surpass the 60% threshold. It should be noted that these thresholds established by Hestenes and Halloun were from semester or even year-long courses, as opposed to two weeks of learning between pre- and post-test for the IPTIR program. Considering this small window of time, the IPTIR participants experienced significant increases in conceptual understanding.

Qualitatively, all of the participants surveyed agreed their participation in the IPTIR program increased awareness or updated knowledge of physics/physical science concepts and provided additional strategies, approaches and resources for providing high-quality instruction in addition to others. Additionally, 100% of the participants surveyed reported their participation led to increased confidence in their ability to be an effective physics/physical science teacher.

Correlations between physics teaching experience and performance on the FCI and TUG-K cannot be definitively stated, because of the participants that completed all three years of the program only two had greater than 5 years of physics teaching experience. Within the population of less than 5 years of physics teaching experience, experience did not have an impact on conceptual assessment pre-test or post-test scores. The two outliers in terms of physics teaching experience had 15 and 20 years of experience, and both scored 100% on both the pre-and post-tests on the FCI and TUG-K. With no participants having between 6 and 15 years of physics teaching experience, it can be hypothesized that increased teaching experience over 5 years correlates with higher content knowledge, but this conclusion cannot be drawn from the presented data.

Correlations among the conceptual assessments are as expected, in that teachers who performed relatively well on any one of the post-tests generally also performed relatively well on the other two assessments. This implies the expected connection between knowledge of Newtonian concepts, the ability to represent knowledge graphically, and scientific reasoning.

Changes in Teacher Pedagogy

Analysis of the participants' teaching pedagogy revealed positive changes consistent with IPTIR's interactive teaching methods. Participants' self-reported classroom practices on average clearly demonstrate a shift away from traditional pedagogy where information is provided by the teacher, towards practices in which teaching for understanding is emphasized through engaging, reflective and authentic learning experiences.

The self-reported changes are supported by standardized videotape analysis of participants' reformed teaching practices. Average participant scores of the video analysis indicated an increasing shift in reformed teaching, consistent with interactive engagement and student-centered techniques. In particular, participants had the greatest improvements in indicators related to lesson design and implementation, procedural knowledge, and student/teacher interactions with 27.2%, 31.2% and 27.9% respective increases in average scores as assessed by the Reformed Teaching Observation Protocol. Each of these three categories are driven by the structure of a lesson, whereas the remaining two indicators, propositional knowledge and communicative interactions, refer to behaviors inherent to the teacher. Effective lesson design and implementation exhibits characteristics such as addressing student misconceptions and providing exploration experiences (MacIsaac & Falconer, 2002). Effective procedural knowledge is indicated by lessons that use a variety of representations and scientific reasoning. Effective student/teacher relationships are indicated by interactive lessons driven by the students and facilitated by the teacher. The program's extensive use of PRISMS PLUS learning cycles and Modeling instruction provided the teachers with lessons that were structured to maximize each of these three indicators, so it would be expected that these areas would have the greatest improvements. It can be predicted that both the propositional knowledge, which refers to knowledge of fundamental physics concepts, and communicative interactions, involving fostering student discourse, will have a delayed impact, possibly improving with continued experience with the physics content and student-centered pedagogy.

No correlations were found between pedagogical shifts and teacher content knowledge, scientific reasoning skills, general teaching experience or physics teaching experience. Therefore it can be concluded that the program is capable of effecting pedagogical shifts towards student-centered techniques for teachers regardless of their teaching experience or knowledge when starting the program.

There was a slight negative correlation between teaching experience and teacher-centered characteristics in the beginning of the program. This can possibly be due to more inexperienced teachers having less practice and therefore less confidence teaching, which may lead to a reliance on more traditional methods and focusing on the textbook. Halim and Meerah's research (2002) proposed that teachers with low-level content knowledge have difficulties forming analogies for instruction, which can often create student misconceptions. For this reason, newer teachers may find comfort in the more scripted nature of teacher-centered instruction, and be apprehensive in an atmosphere that encourages students to ask probing questions. Additionally, data showed that teachers with higher frequencies of teacher-centered characteristics in the beginning of the program had greater changes in student-centered characteristics, which is logical in the sense that there was greater capacity for change towards reformed methods. Teaching experience did not appear impact pedagogical shifts towards student-centered techniques, demonstrating that teachers can make pedagogical shifts regardless of experience or typical classroom practices at the start of the program. The three-year length of the IPTIR program provided extended opportunities for guidance and mentoring, which may have attributed to such positive changes regardless of initial teaching pedagogy or experience.

Changes in Student Achievement

Teaching effectiveness of the IPTIR participants was measured by administering conceptual assessments to their students as pre- and post-tests during the beginning and end of each academic year of the program. The data show that on average each cohort of new students had progressively higher post-test scores, and the student gains increased from year to year for almost all conceptual assessments. These results indicate the participants' students are learning more with each year of the program, a strong indicator that the teacher participants are becoming more effective teachers.

Student conceptual assessment results collected from the IPTIR program support Hake's study of interactive-engagement (IE) vs traditional methods (1998), as IPTIR's average FCI student gain in Year 1 was 0.20, which would fall under Hake's definition of "low-gain" ($g < 0.3$), and were similar to the average gain of traditional courses in Hake's study of $0.23 \pm 0.04$. In Years 2 and 3, IPTIR's average student FCI gain was 0.40, which approaches Hake's average gain in interactive-engagement courses of $0.48 \pm 0.14$. The IPTIR average FCI student gain in Years 2 and 3 falls above Hake's designation of medium gain of 0.30, indicating successful application of interactive engagement instructional methods (Hake, 1998). However it should be noted that no direct correlations were found between individual teacher pedagogical characteristics and student post-test scores nor gains on any of the conceptual assessments.

Correlations show that teachers with higher scores on the FCI post-test were more likely to have students achieve higher scores on their FCI post-test. Physics teaching experience also appeared to impact student FCI post-test scores as well as overall student FCI gain. This supports the previous tentative correlation between the teachers' FCI post-test scores and physics teaching experience.

Student achievement did not correlate with the average high-needs status of the teacher participants' schools, with high-needs defined as at least 30% of K-12 students eligible for free and reduced lunch based on total Title V enrollment (Iowa DoE, 2012d). Schools of 21 of the 35 participants qualified for high-needs status during at least one year of the program, but this status did not appear to impact student test scores or gains.

It can be concluded that in terms of impacting student learning, teacher content knowledge and physics teaching experience appeared to have the largest impact on student learning, but there was also an overall trend of increasing student learning as teacher pedagogy shifted from traditional to increasingly student-centered practices.

Evidence of Success: More High-Quality Physics Teachers

At the outset of the IPTIR program, many of the participants were teaching physics/physical science without having completed the requirements for the State of Iowa Grades 5-12 physics teaching endorsement. As a result of this program 20 out of the 21 teachers that started the program without a completed physics teaching endorsement had acquired this by the end of the program. IPTIR was highly successful in terms of contributing to the number of high school physics teachers in Iowa, an accomplishment that is an important step in addressing the critical need for qualified high school physics teachers in Iowa.

## Limitations

Certain aspects of this study limited interpretation of the results, and these limitations may also serve as suggestions for improvement for further studies.

1. This study focused only on teachers that participated in all three years of the program, which resulted in poor distribution in regards to participant physics teaching experience. Connections made between physics teaching experience are therefore only speculative, as there lacked data from teachers with 6-15 years of physics teaching experience.
2. There was no conceptual assessment data from participants' students prior to the program, which would have provided a baseline measure of participant effectiveness. This lack of data prevented comparison of teacher impact on

students prior to participating in the program to student impact after the program.

4. Conceptual assessments were administered to the teachers at the beginning and end of the first summer's two-week session, which does not capture the full change in teacher content knowledge over the three years of the program.
5. Evaluation of teacher pedagogy was either self-reported or via an instrument that assessed only a single lesson, which does not fully capture a complete picture of a teacher's pedagogical methods. As with many qualitative studies, evaluation of teaching pedagogy was limited by the possibility of bias from both methods of pedagogical assessment.

## Significance and Implications

The results of this study provide evidence of the successful implementation of reformed, interactive instructional methods in a professional development for high school physics teachers. The conclusions drawn from this study lead to the following implications for reformed teaching and professional development:

1. The PRISMS PLUS learning cycles and Modeling Instruction used in the program led to improved teacher content knowledge, and by learning content in a method that can then be transferred to the classroom, teaching pedagogy of the participants also shifted towards more student-centered characteristics. In addition to the implementation of interactive methods in the program, the teacher participants' application of these methods in their classrooms appeared to consequently improve the learning of their students. These measures of

effectiveness imply a need for further professional development opportunities such as the IPTIR program, in order to continually increase the number of qualified high school physics teachers.

2. Conceptual assessment results of the teachers as learners in the program as well as the student conceptual assessment data both support significant conceptual gains for learners in interactive classroom environments. Student data showing increasing pre- to post-test gain in conceptual knowledge over the three years, coupled with teacher pedagogy on average shifting towards more student-centered methods, supports previous studies' claims that interactive engagement teaching methods result in greater student learning (Hake, 1998). The graduated improvements with each year imply that student achievement would continue to increase as the teachers gain more experience with interactive methods. Results of this study contribute to research supporting the nationwide reform movement towards interactive, student-centered learning environments supported by long-term professional development.
3. The program appeared to equally impact participant teaching pedagogy regardless of experience as a science teacher or specifically as a physics teacher. Additionally, data revealed that teachers with a higher frequency of teacher-centered characteristics had greater changes towards reformed methods. This implies that professional development programs that focus on increasing interactive teaching pedagogy can be provided to teachers of all experience levels

and teaching characteristics, and both new and seasoned teachers can learn to incorporate interactive engagement techniques into their classrooms.

4. Data from the study revealed that the more inexperienced teachers started the program with a slightly higher average frequency of traditional teaching characteristics, but these inexperienced teachers were equally as capable of learning and implementing interactive pedagogy as were the more experienced teachers. Newer teachers may be either unfamiliar with interactive techniques, or lack experience implementing these techniques and therefore may rely on the more predictable and comfortable traditional techniques. This has important implications for pre-service physics teacher training. Many pre-service teachers learn physics in traditional college classrooms, but this study shows that pedagogical content knowledge is very valuable to teachers. To produce teachers capable of creating interactive classrooms, teacher preparation should include coursework which builds content knowledge while utilizing interactive engagement techniques. Giving pre-service teachers experiences to build pedagogical content knowledge could build confidence to answer student questions promoted in inquiry settings. The University of Northern Iowa (UNI), in addition to content-specific teaching methods courses required by science education majors, is now offering a course to address this need. Taught by UNI Physics faculty, Resources for Teaching Physics introduces students interactive engagement techniques used in the high school physics classroom by focusing on PRISMS PLUS learning cycles and Modeling Instruction methods.

A need for pre-service pedagogical content knowledge is not limited to the area of physics, or even the field of science. Pre-service teachers in other fields could also benefit from learning content using methods that could then be translated into the classroom.

5. This study found that the most significant correlations with student learning were with teacher content knowledge and physics teaching experience, and the frequency of student-centered practices did not appear to directly impact student conceptual assessment scores. While this may initially appear to contradict previous implications of the benefits of interactive engagement techniques, a lack of direct correlation does not necessarily imply causation, or lack thereof. As Halim and Meerah (2002) proposed, teachers without a solid foundation of conceptual understanding are more likely to be unaware of student misconceptions, and the results from the IPTIR program imply that teacher content knowledge and experience are a major factor in teaching effectiveness. Because student-centered learning requires addressing misconceptions and probing questions, this study supports the implication that deep conceptual understanding must accompany the effective implementation of interactive techniques.

## Recommendations for Future Programs

The need for more professional development programs for high school physics teachers is evident, not only to provide a route for out-of-field teachers to gain physics teaching endorsements, but also to improve the content knowledge of current physics teachers while developing reformed pedagogical techniques. The evidence-based conclusions of this study have led to the following recommendations for future programs:

1. Evidence indicates that both novice and experienced teachers benefit from the program, as well as teachers both strong and weak in content knowledge. Because content knowledge appears to be a limiting factor in the implementation of student-centered techniques, future programs may consider isolating the cohort into participants with weaker or stronger content knowledge. A cohort of participants with weaker content knowledge would include a greater emphasis on learning the content and/or move through the content at a slower pace.
2. To fully capture the impact of the program on teacher learning, it would be recommended to administer conceptual assessments throughout the program as opposed to only at the beginning and end of the initial summer session.
3. To fully capture impact on student learning, it would be recommended to have participants administer conceptual assessments to students at the end of the academic year prior to participating in the program. This would provide data to compare student performance prior to a teacher's participation in the program to student performance throughout a teacher's participation in the program.

4. The three-year aspect of the IPTIR program revealed increasing gains in student learning over the three years of the program, demonstrating increasing teacher effectiveness with continued participation in the program. This data supports recommendations of long-term professional development from Kennedy (1998), Matsler (2010) and Horizon Research, Inc. (Banilower et al., 2013), and the Iowa DoE (2013), emphasizing that future programs should maintain this multiple-year model of professional development.

The data from IPTIR, a three-year program with two-week summer sessions, can be compared to that of the University of Northern Iowa Physics Institute (UNI-PI), a two year program with four-week summer sessions (Escalada & Moeller, 2006). FCI student post-test scores were higher from teachers in the UNI-PI (Year 1: 56.9%, Year 2: 51.8%) than any of the FCI student post-test scores from IPTIR (Year 1: 41.5%, Year 2: 47.7%, Year 3: 48.1%). However the three year IPTIR program resulted in continual increases in student achievement, where this trend was not seen in the UNI-PI data. Additionally, IPTIR student FCI gain continually increased with each year, whereas this was not seen in the UNI-PI data. This does not imply a less effective UNI-PI program, rather it provides evidence that IPTIR's additional year of mentoring appears to be more beneficial to the teacher participants than providing longer summer workshops over only two years.

## Further Research

Further analysis of IPTIR program data or of future programs may lead to increased understanding of the factors that contribute to both teacher and student learning. The following is suggested for further research:

1. Deeper research and comparisons may be performed on data from participants that did not complete all three years of the program to determine benefits and/or differences resulting from one, two and three years of participation.
2. Further correlations may be investigated of student learning with additional participant characteristics, such as teacher confidence, school support, class size, number of class preps, and/or class duration.
3. A long-term study on the cohort of teachers in the IPTIR program may determine if pedagogy continues to shift towards reformed teaching methods after the conclusion of the program, or if participants shifted back towards traditional methods.
4. To determine if reformed teaching methods result in deeper and more permanent understanding, a long term study could be performed on a sampling of participants' students both before and after interactive engagement techniques were implemented in a teacher's classroom. Conceptual assessments could be administered to students that were in a classroom when the teacher more frequently utilized traditional techniques, as well as to students in the same teacher's class after implementing reformed teaching methods. A long-term study

would reveal knowledge retention of students in traditional vs. interactive classrooms.

5. In addition to student achievement data, for future programs pre- and post-surveys could be administered to students determine any changes in student interest in science as a result of changes in classroom atmosphere towards increasingly interactive environments. Data could potentially support findings by Matsler (2010) and the Physics Teaching Resource Agents professional development programs which contained documented increases in interest in science in concert with shifts towards more student-centered classrooms.
6. Future investigations can be performed on teacher participants that teach courses other than physics to determine if the pedagogical skills learned from PRISMS PLUS learning cycles and Modeling Instruction can be transferred to other subject areas. Separate follow-up surveys of classroom practices could be administered to the teachers to complete for multiple courses to determine if the interactive methods continued to develop in their physics courses, and if these methods translated to other subjects.

# REFERENCES

American Association of Physics Teachers. (2009). *Education, qualifications and professional development of secondary physics teachers*. College Park, MD.

American Institute of Physics Statistical Research Center, Tesfaye, C.L. & White, S. (2012). *High School Physics Teacher Preparation*. Retrieved from https://www.aip.org/sites/default/files/statistics/highschool/hs-teacherprep-09.pdf

American Modeling Teachers Association. (n.d.). Retrieved from www.modelinginstruction.org

Balka, D., Hull, T., & Harbin Miles, R. (2013). *What is conceptual understanding?* Retrived from http://www.mathleadership.com/sitebuildercontent/sitebuilderfiles/conceptualUnderstanding.pdf

Banilower, E. R., Smith, P. S., Weiss, I. R., Malzahn, K. A., Campbell, K. M., & Weis, A. M. (2013). *Report of the 2012 national survey of science and mathematics education.* Chapel Hill, NC: Horizon Research, Inc.

Beichner, R. J. (1994). Testing student interpretation of kinematics graphs. *American Journal of Physics, 62*(8), 750-762.

Bucher, A. M. (2009). *A survey of instruments to assess teacher content knowledge in science.* (Master's Thesis). Bowling Green State University, Bowling Green, OH.

Bybee, R. W., Taylor, J. A., Gardner, A., Van Scotter, P., Carlson Powell, J., Westbrook, A., & Landes, N. (2006). *The BSCS 5E instructional model: Origins and effectiveness.* Colorado Springs, CO: BSCS.

Coletta, V. P., & Phillips, J. A. (2005). Interpreting FCI scores: Normalized gain, pre-instruction scores, and scientific reasoning ability. *American Journal of Physics, 73*, 1172.

Cooney, T. M., Escalada, L.T., & Unruh, R.D. (2008). *PRISMS (Physics Resources and Instructional Strategies for Motivating Students) PLUS*. Cedar Falls, IA: University of Northern Iowa Physic Department.

Escalada, L. T., & Moeller, J. K. (2006). *The challenges of designing and implementing effective professional development for out-of-field high school physics teachers.* AIP Conference Proceedings, *818*(1), 11-14.

Escalada, L. T., Morgan, J.M., & Stone, J. (2010). *Iowa Physics Teacher Instruction and Resources (IPTIR) Project: Annual Report*. Cedar Falls, IA.

Escalada, L. & Stone, J. (2009, June 8). Teacher Survey of Classroom Practices [Survey], as adapted from the CCSO/State Collaborative on Assessment and Student Standards. (Original survey published January 14, 1997.

Fisch, K. & McCloud, S. (2012). Did you know?/Shift happens. Retrieved Dec. 1, 2012 from http://www.youtube.com/watch?v=ljbI-363A2Q

Goris, T. V., & Dyrenfurth, M. J. (2010). *Students' misconceptions in science, technology and engineering.* ASEE Illinois/Indiana Section Conference proceeding, Purdue University, West Lafayette, IN.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64-74.

Halim, L., & Meerah, S. M. M. (2002). Science trainee teachers' pedagogical content knowledge and its influence on physics teaching. *Research in Science & Technological Education*, *20*(2), 215-225.

Hestenes, D. (1987). Toward a modeling theory of physics instruction. *American Journal of Physics, 55*(5), 440-454.

Hestenes, D. (1996). Modeling methodology for physics teachers. *Proceedings of the International conference on Undergraduate Physics Education.* College Park, MD.

Hestenes, D. & Halloun, I. (1995). Interpreting the force concept inventory: a response to Huffman and Heller. *Physics Teacher, 33*, 502–506.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, *30*(3), 141-158.

Huffman, D., & Heller, P. (1995). What Does the Force Concept Inventory Actually Measure? *The Physics Teacher*, *33*(3), 138-43.

Iowa Department of Education. (2012a). *Iowa Core, Physical Science*. Retrieved from https://www.educateiowa.gov/pk-12/iowa-core/science/science-%C2%BB-physical-science

Iowa Department of Education. (2012b). *Teacher Shortage Areas*. Retrieved from https://www.educateiowa.gov/teacher-shortage-areas

Iowa Department of Education. (2012c). *Title II - Part A - Teacher and Principal Training and Recruiting*. Retrieved from https://www.educateiowa.gov/pk-12/title-programs/title-ii/title-ii-part-teacher-and-principal-training-and-recruiting

Iowa Department of Education. (2012d). *Title V- Part A – Innovative Programs*. Retrieved from https://www.educateiowa.gov/pk-12/title-programs/title-v-part-innovative-programs

Iowa Department of Education. (2013). *Iowa Professional Development Model*. Retrieved from https://www.educateiowa.gov/pk-12/educator-quality/iowa-profesional-development-model

Jackson, J., Dukerich, L., & Hestenes, D. (2008). Modeling Instruction: An Effective Model for Science Education. *Science Educator*, *17*(1).

Karplus, R. & Butts, D. P. (1977). Science teaching and the development of reasoning. *Journal of Research in Science Teaching*, *14*, 169–175.

Keeley, P. (Ed.). (2008). *Science formative assessment: 75 practical strategies for linking assessment, instruction, and learning.* Arlington, VA. SAGE.

Kennedy, M. (1998). Form and substance in in-service teacher education. Research Monograph. *Annual meeting of the American Educational Research Association*. San Diego, CA.

Knight, R. (2004). *Five easy lessons: strategies for successful physics teaching.* San Francisco, CA: Addison Wesley.

Lawson, A. (1978). The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching, 15*(1), 11-24.

Lawson, A., Abraham, M., & Renner, J. (1989). A theory of instruction: Using the learning cycle to teach science concepts and thinking skills. *Monographs of the National Association for Research in Science Teaching*, 1, 1-57.

Loughran, J., Mulhall, P., & Berry, A. (2008). Exploring pedagogical content knowledge in science teacher education. *International Journal of Science Education, 30*(10), 1301-1320.

MacIsaac, D., & Falconer, K. (2002). Reforming physics instruction via RTOP.*The Physics Teacher*, *40*(8), 479-485.

Maier, S.J. & Marek, E.A. (2006). The learning cycle: A reintroduction. *The Physics Teacher,* (44) 2: 109–13.

Marx, J.D. &Cummings, K. (2007). Normalized change. *American Journal of Physics, 75*, 87-91.

Matsler, K.J. (2010). *Final report AAPT/PTRA Rural Project*. Retrieved from http://www.aapt.org/PTRA/upload/2010-NSF-Final-Report.pdf

McDermott, L.C. (1991). What we teach and what is learned- closing the gap. *American Journal of Physics, 59*(4), 301-315.

Meltzer, D. E. (2005). Relationship between students' problem-solving performance and representational format. *American Journal of Physics*, *73*, 463.

Minstrell, J. (2000). Student thinking and related assessment: Creating a facet-based learning environment. *Grading the nation's report card: Research from the evaluation of NAEP*, 44-73. Washington D.C. National Academy Press.

National Institutes of Health, Office of Science Education. (n.d.). Retrieved from http://science.education.nih.gov/SciEdNation.nsf/EducationToday1.html

National Science Teachers Association (2013). Position Statement: Scientific Inquiry. Retrieved from http://www.nsta.org/about/positions/inquiry.aspx

Next Generation Science Standards. (2013). Next Generation Science Standards. Retrieved from http://www.nextgenscience.org/next-generation-science-standards

Park, S., Jang, J.-Y., Chen, Y.-C., & Jung, J. (2010). Is Pedagogical Content Knowledge (PCK) Necessary for Reformed Science Teaching?: Evidence from an Empirical Study. *Research in Science Education, 41*(2), 245-260.

Piburn, M. & Sawada, D. (n.d.). *Reformed teaching observation protocol (RTOP) reference manual* (ACEPT Technical Report No. IN00-3). Phoenix: Arizona State University, Arizona Collaborative for Excellence in the Preparation of Teachers.

Radford, D. L. (1998). Transferring theory into practice: A model for professional development for science education reform. *Journal of Research in Science Teaching, 35*(1), 73-88.

Sawada, D. & Piburn, M.(2000). *Reformed Teaching Observation Protocol (RTOP) Training Guide* (ACEPT Technical Report No. IN00-2). Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Teachers.

Sawada, D., Piburn, M., Falconer, K., Turley, J., Benford, R., & Bloom, I. (2000). *Reformed Teaching Observation Protocol (RTOP)* (ACEPT Technical Report No. IN00-1). Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Teachers.

Sahin, E., & Yağbasin, R. (2012). Determining which introductory physics topics pre-service physics teachers have difficulty understanding and what accounts for these difficulties. *European Journal of Physics, 33*(2), 315.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*(1), 1-23.

Shulman, L. S. (1994). Those who understand: Knowledge growth in teaching. *Teaching and Learning in the Secondary School,* 125-133.

Staver, J. R. (1998), Constructivism: Sound theory for explicating the practice of science and science teaching. *Journal of Research in Science Teaching, 35*, 501–520.

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

Task Force on Teacher Education in Physics. (2012). *Transforming the Preparation of Physics Teachers:A Call to Action*. American Physical Society, College Park, MD.

Tesfaye, C.L., & White, S. (2011). High school physics teacher preparation: Results from the 2008-09 Nationwide Survey of High School Physics Teachers. Focus On. *Statistical Research Center of the American Institute of Physics*. College Park, MD.

Ucar, S. (2012). How Do Pre-Service Science Teachers' Views on Science, Scientists, and Science Teaching Change Over Time in a Science Teacher Training Program? *Journal of Science Education and Technology*, *21*(2), 255-266.

U.S. Department of Education, National Center for Education Statistics (2011-2012). *Schools and Staffing Survey*, "Public School Data File." Retrieved from http://nces.ed.gov/surveys/ruraled/tables/c.1.c.-1_1112.asp?refer=urban

van Driel, J. H., Jong, O. D., & Verloop, N. (2002). The development of preservice chemistry teachers' pedagogical content knowledge. *Science Education, 86*(4), 572-590.

van Driel, J. H., Verloop, N., & de Vos, W. (1998). Developing science teachers' pedagogical content knowledge. *Journal of Research in Science Teaching, 35*(6), 673-695.

Wells, M., Hestenes, D., & Swackhamer, G. (1995). A modeling method for high school physics instruction. *American Journal of Physics*, *63*(7), 606-619.

Wenning, C.J. (2005). Whiteboarding and Socratic dialogues: questions and answers. *Journal of Physics Teacher Education Online, 3*(1), 3-10.

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education, 92*, 941-967.

Yerrick, R., Parke, H., & Nugent, J. (1997). Struggling to promote deeply rooted change: The "filtering effect" of teachers' beliefs on understanding transformational views of teaching science. *Science Education*, *81*(2), 137-159.

# APPENDIX A

## PARTICIPANT DEMOGRAPHICS

| Partici-pant # | Year(s) participated | Years Qualified as High Needs Status | AEA | School Student population | Subject area expertise | Other subject areas |
|---|---|---|---|---|---|---|
| 1 | 2009, 2010, 2011 | NA | 267 | 165 (9-12) | General Science | Ecology |
| 2 | 2009, 2010, 2011 | 2009-2010 (30.3%) | Mississippi Bend | 321 (9-12) | Biology, Chemistry | Earth Science, Physics |
| 3 | 2009, 2010, 2011 | Unknown | 267 | 215 (9-12) | Math, Science | Algebra, Chemistry, Physics |
| 4 | 2009, 2010 | 2007-2008 (48.6%) 2008-2009 (45.9%) 2009-2010 (47.2%) | Heartland | 475 (9-12) | All science | Physics, Chemistry, Astronomy |
| 5 | 2010, 2011* | NA | Grant Wood | 800 (9-12) | Biology, Chemistry | Physics |
| 6 | 2009, 2010, 2011 | NA | Great Prairie | 195 (7-12) | Biology, Chemistry | Physical Science, Physics |
| 7 | 2011 | None | Grant Wood | 128 (9-12) | Biology | Chemistry, Physics |
| 8 | 2009, 2010, 2011 | 2007-2008 (57.1%) 2008-2009 (59.3%) 2009-2010 (67.1%) | Prairie Lakes | 640 (9-12) | Physical Science, Biotech | Earth Science |

Table continues

| Partici-pant # | Year(s) participated | Years Qualified as High Needs Status | AEA | School Student population | Subject area expertise | Other subject areas |
|---|---|---|---|---|---|---|
| 9 | 2010, 2011 | None | Keystone | 600 (9-12) | Physics, Math | Physical Science, Environmental Science |
| 10 | 2009, 2010, 2011* | 2007-2008 (37.5%) 2008-2009 (40.1%) 2009-2010 (43.6%) | 267 | 1200 (9-12) | Math, Physical Science | Physics of Technology |
| 11 | 2009, 2010 | 2008-2009 (30.7%) 2009-2010 (31.2%) | 267 | 88 (9-12) | Biology, Earth Science | Physical Science, Chemistry |
| 12 | 2009, 2010, 2011 | None | Northwest | 360 (K-12) | Biology, Physics, Chemistry | NA |
| 13 | 2010, 2011 | 2007-2008 (38.0%) 2008-2009 (44.0%) 2009-2010 (45.9%) | Great Prairie | 265 (9-12) | Chemistry | Physical Science, Physics |
| 14 | 2009 | 2007-2008 (57.1%) 2008-2009 (59.3%) 2009-2010 (67.1%) | Prairie Lakes | 770 (9-12) | Mechanics | Physics, Chemistry |
| 15 | 2009, 2010, 2011 | 2009-2010 (30.4%) | Grant Wood | 1,500 (9-12) | All science | Physics, Foundations Of Science |

Table continues

| Partici-pant # | Year(s) participated | Years Qualified as High Needs Status | AEA | School Student population | Subject area expertise | Other subject areas |
|---|---|---|---|---|---|---|
| 16 | 2009, 2010 | None | Grant Wood | 412 (K-12) | Chemistry | Physical Science |
| 17 | 2010, 2011* | NA | Loess Hills | 350 (9-12) | Physics, Chemistry | NA |
| 18 | 2010 | 2007-2008 (58.1%)<br>2008-2009 (59.6%)<br>2009-2010 (63.4%) | 267 | 1,546<br>(K-12) | Chemistry | NA |
| 19 | 2010, 2011 | 2008-2009 (30.6%)<br>2009-2010 (34.53%) | 267 | 575 (9-12) | Physical Science, Ag | Physics |
| 20 | 2009, 2010, 2011 | None | Heartland | 650 (9-12) | Biology, Chemistry | Physics |
| 21 | 2009, 2010, 2011 | 2007-2008 (31.7%)<br>2009-2010 (31.7%) | Keystone | 267 (9-12) | Earth Science, Chemistry | Physics, Physical Science |
| 22 | 2009, 2010, 2011 | None | 267 | 165 (9-12) | Biology, Chemistry | Physical Science |
| 23 | 2009, 2010, 2011 | 2007-2008 (50.0%)<br>2008-2009 (49.2%)<br>2009-2010 (56.3%) | Northwest | 1,400 (9-12) | Physics, Chemistry | |

Table continues

| Partici-pant # | Year(s) participated | Years Qualified as High Needs Status | AEA | School Student population | Subject area expertise | Other subject areas |
| --- | --- | --- | --- | --- | --- | --- |
| 24 | 2009, 2010, 2011 | 2007-2008 (34.2%)<br>2008-2009 (36.5%)<br>2009-2010 (41.4%) | Prairie Lakes | 193 (9-12) | Physics, Chemistry | |
| 25 | 2009, 2010, 2011 | 2007-2008 (42.7%)<br>2008-2009 (40.5%)<br>2009-2010 (43.5%) | Prairie Lakes | 160 (9-12) | Biology, Chemistry | Physical Science |
| 26 | 2011 | None | Mississippi Bend | 1,177 (9-12) | Biology, Chemistry | |
| 27 | 2011 | 2007-2008 (30.3%)<br>2008-2009 (37.0%)<br>2009-2010 (37.2%) | Prairie Lakes | 240 (9-12) | Biology | Physical Science |
| 28 | 2009, 2010, 2011 | 2007-2008 (40.7%)<br>2008-2009 (41.5%)<br>2009-2010 (49.3%) | Green Valley | 348 (9-12) | Biology, Chemistry | Physical Science |
| 29 | 2009, 2010, 2011 | 2007-2008 (30.3%)<br>2008-2009 (37.0%)<br>2009-2010 (37.2%) | Prairie Lakes | 714 (K-12) | Biology, Anatomy & Physiology | Physics, Chemistry |

Table continues

| Partici-pant # | Year(s) participated | Years Qualified as High Needs Status | AEA | School Student population | Subject area expertise | Other subject areas |
|---|---|---|---|---|---|---|
| 29 | 2009, 2010, 2011 | 2007-2008 (30.3%)<br>2008-2009 (37.0%)<br>2009-2010 (37.2%) | Prairie Lakes | 714 (K-12) | Biology, Anatomy & Physiology | Physics, Chemistry |
| 30 | 2009, 2010 | None | Prairie Lakes | 1250 (K-12) | Biology | Physics, Chemistry |
| 31 | 2010, 2011 | 2007-2008 (35.7%)<br>2008-2009 (38.4%)<br>2009-2010 (41.6%) | Great Prairie | 700 (9-12) | Biology, Chemistry | NA |
| 32 | 2009, 2010, 2011 | None | Heartland | 140 (9-12) | Chemistry | Physics, Physical Science |
| 33 | 2009, 2010 | 2007-2008 (34.3%)<br>2008-2009 (35.8%)<br>2009-2010 (45.5%) | 267 | 291 (7-12) | Biology, Earth Science, Physical Science | NA |
| 34 | 2010, 2011 | 2007-2008 (50.0%)<br>2008-2009 (49.2%)<br>2009-2010 (56.3%) | Northwest | 1,350 (9-12) | Physics, Biology, Chemistry | NA |
| 35 | 2009, 2010, 2011 | None | Grant Wood | 1654 (K-12) | Biology | Physics, Earth Science |

*Summer only