A model for assessing the effectiveness of professional development for improving student learning

John Francis O'Connell

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Curriculum and Instruction Commons

Recommended Citation
https://lib.dr.iastate.edu/etd/10677

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
A model for assessing the effectiveness of professional development for improving student learning

by

John Francis O’Connell

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Education

Program of Study Committee:
Gary Phye, Co-Major Professor
Carl Smith, Co-Major Professor
Tom Andre
Dale Niederhauser
Dan Russell
Ann Thompson

Iowa State University
Ames, Iowa

2009

Copyright © John Francis O’Connell, 2009. All rights reserved.
Table of Contents

LIST OF FIGURES ......................................................................................................................... iv

LIST OF TABLES ............................................................................................................................ v

ABSTRACT ..................................................................................................................................... vi

CHAPTER 1. INTRODUCTION ..................................................................................................... 1

Statement of the Problem ............................................................................................................. 1

Historical Perspective .................................................................................................................. 2

Delivery of Professional development ...................................................................................... 4

Rationale for the current study .................................................................................................. 5

Growth Curves ............................................................................................................................ 6

Summary ..................................................................................................................................... 7

CHAPTER 2. REVIEW OF THE LITERATURE ............................................................................. 9

Background .................................................................................................................................. 9

School Reform and the Increased Need for Professional Development ....................... 10

First Wave, 1960-1990 .................................................................................................................. 11

Second Wave, 1990-2000 ............................................................................................................. 12

Third Wave, 2000-2009 ............................................................................................................... 13

Current Status ............................................................................................................................. 14

Students’ Achievement Growth .............................................................................................. 15

Teachers’ Growth ....................................................................................................................... 17

Iowa Professional Development Model (IPDM) ..................................................................... 18

Technology in Support of Professional Development ......................................................... 19

Summary ..................................................................................................................................... 21
<table>
<thead>
<tr>
<th>CHAPTER 3. METHODOLOGY</th>
<th>22</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose of the Study</td>
<td>22</td>
</tr>
<tr>
<td>Research Hypotheses</td>
<td>22</td>
</tr>
<tr>
<td>Research Design</td>
<td>25</td>
</tr>
<tr>
<td>Achievement Measures</td>
<td>26</td>
</tr>
<tr>
<td>Educational Intervention</td>
<td>27</td>
</tr>
<tr>
<td>Data Collection Design</td>
<td>28</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 4. ANALYSIS OF DATA</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Question 1</td>
<td>30</td>
</tr>
<tr>
<td>Hypothesis I</td>
<td>32</td>
</tr>
<tr>
<td>Hypothesis II</td>
<td>34</td>
</tr>
<tr>
<td>Research Question 2</td>
<td>36</td>
</tr>
<tr>
<td>Hypothesis III</td>
<td>36</td>
</tr>
<tr>
<td>Summary</td>
<td>42</td>
</tr>
</tbody>
</table>

| CHAPTER 5. DISCUSSION         | 43 |
| Empirical Data Model          | 43 |
| Discussion of Research Question 1 and Hypothesis I | 43 |
| Discussion of Research Question 1 and Hypothesis II | 44 |
| Conclusions                   | 48 |
| Shortcomings of the Study     | 48 |
| Future Research               | 49 |

APPENDIX .......................... 52
REFERENCES .......................... 53
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Iowa Professional Development Model</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>Iowa Professional Development Model</td>
<td>19</td>
</tr>
<tr>
<td>3</td>
<td>Comparison of experimental and comparison schools</td>
<td>33</td>
</tr>
<tr>
<td>4</td>
<td>Mathematics and Reading scores for experimental school students</td>
<td>35</td>
</tr>
<tr>
<td>5</td>
<td>Fully recursive model for Heartland data</td>
<td>39</td>
</tr>
<tr>
<td>6</td>
<td>Trimmed model for Heartland mathematics experimental 8th grade students</td>
<td>41</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1: Means for the significant two-way year by subject interaction 35
Table 2: Within-subject ANOVA summary table 36
Table 3: Intercorrelations, Mean, and SD for measured variables 42
ABSTRACT

For most of the last 50 years, teachers and administrators have perceived professional
development and the instructional role primarily in terms of what one individual does with
classes of others. Teachers, like their students, were considered to be rather passive
acceptors of the instruction, rather than active modifiers. Thus, there was very little effort to
gather empirical evidence that professional development worked especially when it came to
its impact on student achievement. This project proposes a research design to gather
empirical data on the Iowa Professional Development Model supported by technology. In
this study educators were provided with an opportunity to integrate or vertically transfer
much of what they learn through their professional development activities. The primary
means of achieving this was through the building of a mature professional learning
community connected by technology providing the key components of coaching and
feedback reducing teacher isolation. These communities led to increased teacher
understanding of the strategies, stimulated change, and resulted in increased student
achievement.
CHAPTER 1. INTRODUCTION

Statement of the Problem

Currently, there is a renaissance of field experimentation in evaluating educational interventions (Shadish & Cook, 2009). In their Annual Review of Psychology chapter, these authors identify current models of field experimentation that are employed to provide credible data of evidence-based practices in schools. Currently, the best research models in field experimentation investigating the effectiveness of educational interventions include randomized experiments, which are the “gold standard” or varieties of quasi-experimental\(^1\) experiments. There is not a single class of quasi-experimental designs that fit all research questions involving the educational system. Rather, if an experimental design cannot be employed for data collection, quasi-experimental models that best address stakeholder research questions is the preferred approach to providing credible data addressing educational policy issues and educational practices at the building level.

This dissertation is an effort to develop a field experimentation model that would provide policy makers and educational professionals in Iowa schools with a method for evaluating the impact of professional development. This field research model is designed to assess the impact of professional development delivered by an Area Education Agency’s (AEA) professional development team to middle school mathematics teachers who

---

\(^1\) Experiment: A study in which an intervention is deliberately introduced to observe its effects. Randomized Experiment: An experiment in which units are assigned to receive the treatment or an alternative condition by a random process such as the toss of a coin or a table of random numbers. Quasi-Experiment: An experiment in which units are not assigned to conditions randomly. (Shadish, Cook, Campbell, 2006, p. 3)
implement instructional strategies in order to promote student growth in math achievement during the eighth grade.

**Historical Perspective**

The reform of public education has been a major focus of policymakers at the local, state, and national levels since the turn of the century. Reform efforts have been sparked by individuals such as John Dewey and others, and the publication of reports like the 1983’s *A Nation at Risk* (National Commission on Excellence in Education, 1983). The latter resulted in the charge that American education was failing the country and its youth. Policymakers citing declining standardized test scores both nationally and internationally, began to demand for increasing rigor and accountability in America’s public schools.

These demands for major educational reforms were based on certain historical assumptions. First, “the belief that a good education leads to individual financial success, higher personal status, and the benefits of a flourishing economy had become largely unquestioned by the late twentieth century” (Cuban, 2005, p.10). A second reason for the broad acceptance of these reforms was the tendency in the United States to regard more and better public education as the solution for many of the nation’s social, economic, and political problems (Cuban, 2005). These assumptions along with the *A Nation at Risk* report led to the development of standards for various curricula. These new curriculum standards require teachers to make changes in what and how they teach and in their roles in classrooms and schools. These changes required an investment in the teachers’ professional development.

According to Noyce (2006, p.36), “we must acknowledge first that the reason we do professional development is so that students will learn more. Every other outcome is
important only insofar as it leads to that.” The primary premise for this statement is that teachers who know and can do more effective teaching in the classroom help students learn more. Teaching teachers appeals to our intuition as a high-leverage strategy for boosting student achievement. Yet there is a lack of empirical data to support this notion. Guskey helps frame the discussion by saying, “In particular, policy makers and educational leaders want specific evidence of the impact of professional development activities on well-defined student learning outcomes” (Guskey, 2002a, p. 45). Those responsible for professional development have generally assumed a strong and direct relationship between professional development for educators and improvements in student learning. “Few have been able to describe the precise nature of that relationship” (Guskey & Sparks, 2002, p.1).

This study attempts to provide a research design that is based on empirical evidence that helps address this lack of evidence regarding the impact of professional development on student achievement. The questions addressed in this study are; 1) do professional development activities conducted within the context of the Iowa Professional Development Model (IPDM) improve student achievement and 2) does the use of technology to create professional learning communities support the teachers’ implementation of the professional development activities?

The study’s focus area was in the content area of eighth grade mathematics at middle schools located in central Iowa (Heartland Area Education Agency). One of the outcomes of this study is to suggest a model that uses credible data to provide empirical evidence that professional development activities do improve student learning. This is not to say that there is a cause-effect relationship between the two. Teaching performance is only one of many factors that predict student learning. Examples of some of the other factors that have been
shown to impact student learning are: expenditures per student, quality of facilities, student-to-teacher ratio, student mobility socio-economic status, etc.

**Delivery of Professional development**

Iowa’s response to the question of effective delivery of professional development activities is the Iowa Professional Development model (Iowa Department of Education, 2002). This model is based on suggestions by Joyce and Showers (1982). The IPDM shown in Figure 1, requires a continuous cycle of formative and summative assessment of 1) training and learning opportunities for teachers, and 2) evidence of collaborative implementation of professional development activities. Thus, successful reform is in part predicated on the development of a “learning community” of teachers within a school district as a part of the districts comprehensive school improvement plan. Finally, successful school improvement is defined as growth in student learning and achievement.

**Figure 1. Iowa Professional Development Model**
The IPDM was designed to break with the past school environment in which teachers learned to work alone, unengaged for the most part in group decisions or the necessity to coordinate activities with others (Wagner, 2001). The IPDM supported by technology would provide educators with an opportunity to interact with each other both within and outside of their school building ending their isolation. It would also help them to integrate or vertically transfer much of what they learn through professional development into their routine practice with the support and resources necessary to fully implement it and thereby increase the probability that it will lead to increased student achievement.

To demonstrate that technology-supported professional development works, it is necessary to develop a research design that uses empirical data that produces evidence-based results. The need for empirical data results was pointed out in a review of technology related professional development papers by Davis and Thompson (2005). Reviewing more than 100 articles and reports on evaluations of technology-related professional development, they eliminated the majority of this work because the evaluation described focused upon the lowest level of Guskey’s categories (participant reaction), or did not use carefully constructed/tested instruments or failed to report clear result (Davis & Thompson, 2005). They conclude that “Our review clarifies the need for creating a design that allows for the collection of scientifically based evidence and provides guidance for how to do this” (Davis & Thompson, 2005, p.2).

Rationale for the current study

In order to evaluate the effectiveness of the IPDM, the following logical model was proposed. Professional development as planned and delivered to schools by the Area Education Agency must be implemented in the classroom in order for student achievement
gains to be observed during the academic year. Further, in the present study the program evaluation effort was situated in eighth grade mathematics. Since not all of the elements of the IPDM can be evaluated in a single study, this study focused on the system components of (a) professional development, (b) teacher collaboration/implementation, and (c) student growth in mathematics achievement during the eighth grade. Mathematics achievement in this case was assessed using the mathematics total score from the Iowa Test of Basic Skills (ITBS) for the 2005-06 and the 2006-07 academic school years.

**Growth Curves**

As described, evidence of successful school improvement in Iowa is predicated on the successful demonstration of growth in student achievement resulting from positive classroom learning experiences. Currently, there are two approaches to determining student growth in math achievement. The first approach, the assessment of adequate yearly progress is an approach mandated by the No Child Left Behind Education Act (2002). This approach involves the annual determination of progress demonstrated by non-proficient students in any classroom. The focus is on moving non-proficient students to a level of proficiency defined for each grade by performance on the Iowa Test of Basic Skills\(^2\) (Iowa Consolidated State Application Accountability Workbook for State Grants under Title IX, 2004, p.8). This approach is commendable but unfortunately does not focus on all students within the grade.

---

\(^2\) In Iowa, all public schools and LEAs will be held to the same process and criteria for making adequate yearly progress toward 100% proficiency by the 2013-14 school year. For purposes of AYP accountability, all public schools and LEAs will be judged by performance and improvement on the Iowa Tests of Basic Skills (ITBS) and the Iowa Tests of Educational Development (ITED). These measures are the common comparable measures across all school districts, thus, ensuring fairness, validity, and reliability when making unbiased, rational, and consistent determinations of the annual progress of LEAs and schools within the state. All schools and districts will be expected to make improvement in student achievement. (Iowa Consolidated State Application Accountability Workbook, page 8)
A second approach is one taken in the current study. This is a growth curve approach where a cohort of all students present during the academic year is followed from the beginning of the grade to the end of the grade. During this interval, the educational intervention (implementation by the teacher of professional development content) has an opportunity to facilitate student growth in achievement. In this case, student growth in math achievement as measured by the Iowa Test of Basic Skills (ITBS).

These two approaches are complementary and data from each approach facilitates policy development and analysis for stakeholders at the state department of education. Further, Iowa is one of nine states identified by the federal department of education as engaging in preliminary pilot studies determining the effectiveness of growth curve analysis for reporting growth in student achievement. There is however an advantage to growth curve analysis. The longitudinal cohort approach provides information about effective classroom instruction and curriculum alignment that is not possible using only adequate yearly progress data. Thus, the growth curve analysis model of data collection in field experimentation will provide a basis for the future evaluation of the Iowa Core Curriculum starting in 2012.

Summary

The reform of public education that is based on curriculum standards require teachers to make tremendous changes in what and how they teach and in their roles in classrooms and schools. These changes required an investment in the teachers’ professional development. The reform efforts have led to an increased emphasis on teacher professional development, which has been assumed to result in increased student achievement. As the review of literature will show, there is a strong belief among professional development providers and organizations that professional development does improve student achievement but there is a
lack of empirical data that supports this claim. The purpose of this study was to develop a research design that would provide such data.
CHAPTER 2. REVIEW OF THE LITERATURE

Background

The story line for the literature reviewed in Chapter 2 starts with a brief consideration of teachers’ changing roles in today’s classrooms. This change is due, in part, to the emphasis on educational reform in the United States that can be traced back to the middle of the 1980s and in Iowa to the turn of the twenty-first century.

Today’s classroom environment is in transition as a result of the development of newly defined standards for both curriculum content and student performance (Killion, 2002). Add to this reform in curriculum and student expectations the heavy infusion of technology into the school environment during the past decade, and the need for high quality teaching becomes apparent (Fishman, Marx, Best & Tal, 2003). In essence, this emphasis on curriculum content and student performance is the background for the quality teacher initiative that has taken two forms. The first relates to the preparation of new teachers. The second is the more massive effort and concerns the development of quality teaching skills of current classroom teachers (Hawley & Valli, 2000).

This massive effort requires intense professional development for currently employed staff in teaching positions. This nation-wide professional development effort is being undertaken based on the assumption that the classroom teacher is the most important facilitator of student learning outcomes in the school environment (Farrace, 2002; National Staff Development Council, 2001).

There is a considerable body of professional literature on the effectiveness of professional development. Much of it represents widely agreed upon expert opinion, or
hypotheses, about the effectiveness of particular programs and their link to student learning (Elmore, 2000; Darling-Hammond, 2000; Furhman, 2001; Hattie, 2003). Unfortunately, there is little empirical evidence of a causal relationship between professional development and gains in student achievement. Although the causal relationship would intuitively appear to be logical, there is little empirical data to support the logical conclusion (Flowers & Mertens, 2003; Killion, 2002).

**School Reform and the Increased Need for Professional Development**

In today’s reform environment there is an increasing emphasis on teacher professional development in an effort to improve student achievement. Schools are slowly beginning to shift from what teachers should teach to what students should learn (Killion, 2002). Academic standards emphasize what students *should know and be able to do* rather than the scope and sequence of curricula. These new curriculum standards require teachers to make tremendous changes in what and how they teach and in their perceptions of their roles in classrooms and schools (Killion, 2002).

These academic standards are based on the voluntary national standards that arose after the publication of *A Nation at Risk* (National Commission on Excellence in Education, 1983). The national standards, starting in 1989 with the National Council of Teachers of Mathematics’ (NCTM) standards, envision classrooms as places where students regularly explore interesting problems; for example, using important concepts rather than memorizing isolated facts and procedures (National Council of Teachers of Mathematics, 1989). These *active* classrooms include small group work, work with concrete materials, and problem solving in the context of projects. Students are encouraged to communicate ideas orally and
in writing by questioning procedures and results, discussing and evaluating alternative approaches, and providing written explanations of their reasoning (NCTM, 1989).

Currently, teachers are seen as facilitators of goals rather than the exclusive source of knowledge (Lindquist, Dossey, & Mullis, 1995). This does not mean that teachers’ roles have diminished; indeed, they are more important than ever (Porter, Garet, Desimone, Yoon & Birman, 2000). The assumption is that students cannot learn the knowledge and skills they need to perform up to standard if teachers do not have the knowledge and skills to teach them (Guskey, 2000). Unfortunately, many teachers do not (Killion, 2002). Content standards may call for students to develop knowledge and skills that teachers have not previously addressed. Many teachers, especially teachers in middle schools, are teaching outside of the subject area of their pre-service preparation (Killion, 2002). Other teachers have earned only the minimal credits necessary to qualify for their degree and license to teach (Killion, 2002). In other words, the new emphasis on standards and helping students meet them is just as challenging for teachers as for students.

Changing teacher roles demanded new professional development and research studies on the efficacy of professional development on student achievement (Guskey, 2002a; Killion, 2002, Marzano, 2003). It was this premise that served the Iowa Department of Education as the impetus to develop the Iowa Professional Development Model (IPDM). This is not a new endeavor. Rather, from an historical perspective, three distinct waves of research efforts linking professional development and student achievement can be identified.

**First Wave, 1960-1990**

Research on the links between teacher learning and student achievement can be divided into three waves (Holland, 2005). The first wave began in the 1960s and focused
primarily on *generic* teaching skills, such as allocating class time, providing clear classroom demonstrations, assessing student comprehension during lectures, maintaining attention, and grouping students (Holland, 2005). These studies showed small to moderate positive effects on students’ basic skills, such as phonetic decoding and arithmetic operations; in a few cases, reasoning skills also improved (Holland, 2005). For example, in an experimental study of fourth-grade mathematics in urban schools serving primarily low-income families, student achievement was greater when teachers emphasized active whole-class instruction - giving information, questioning students, and providing feedback - and more frequent reviews, among other measures. Student achievement also was enhanced when teachers learned to follow the presentation of new material with *guided practice*; that is, asking questions and supervising exercises (Holland, 2005).

**Second Wave, 1990-2000**

In the 1990s, a second wave of research delved deeper into student learning, focusing on students’ reasoning and problem solving potentials rather than solely on basic skills (Holland, 2005). Results suggested that professional development can influence teachers’ classroom practices significantly and lead to improved student achievement when it focuses on (a) how students learn particular subject matter, (b) instructional practices that are specifically related to the subject matter and how students understand it, and (c) strengthening teachers’ knowledge of specific subject-matter content. Close alignment of professional development with actual classroom conditions also is key (Holland, 2005).
Third Wave, 2000-2009

The third wave saw an increased emphasis on teacher accountability. The bottom line for professional development was to help teachers become better teachers and improve student learning and report that progress to the public (Belzer, 2003; Elmore, 2002; Farnsworth, Shaha, Bahr, Lewis & Benson, 2002; Guskey, 2002b; Killion, 2002; Lewis & Shaha, 2003). Professional and staff development programs were intended to equip teachers with new or refined skills and techniques for achieving better results for their students and for helping teachers themselves to be more confident and capable (Holland, 2005).

Educational literature in the last decade has argued that what teachers know and do impacts what their students know and do (Puma & Raphael, 2001; Birman, Desimone, Garet, & Porter, 2000). Among educational researchers, a consensus is emerging about particular characteristics of high quality professional development. These characteristics include a focus on content and how students learn content; in-depth, active learning opportunities; links to high standards; opportunities for teachers to engage in leadership roles; and the collective participation of groups of teachers from the same school or grade in the development of learning communities of teachers (Garet, Porter, Desimone, Birman, & Yoon, 2001). Although lists of characteristics such as these commonly appear in the literature on effective professional development, there is little direct evidence on the extent to which these characteristics are related to better teaching and increased student achievement (Hiebert, 1999; Loucks-Horsley, Hewson, Love, & Stiles, 1998; U.S. Department of Education, 1999).
Current Status

Research results on what the model of educational effectiveness and effective teacher professional development should look like is quite mixed. For example, the literature on educational effectiveness suggests a connection between quality teachers, their professional development, and their students’ achievement (Rowe, 2003; Hattie, 2003). More specifically, some studies reported a convergence in the research literature on the characteristics that define effective professional development that are connected with students’ learning (Cohen and Hill, 2000; Thompson, 2003; Ingvarson, Meiers and Beavis, 2005; Kennedy, 1998).

In contrast, according to Guskey’s (2003) review of 13 sources responsible for listing the characteristics of effective professional development, there was not one characteristic that was mentioned on all lists. Guskey’s analysis of the lists of the characteristics of effective professional development yielded three related conclusions. First, there appears to be little agreement among professional development researchers or practitioners regarding the criteria for effectiveness in professional development (Guskey, 2003). Second, perhaps because of the lack of agreement on the criteria for effectiveness, many of the currently identified characteristics of effective professional development can be best described as "Yes, but..." statements. For example, yes, enhancing teachers' content and pedagogic knowledge is important, but existing research is limited mainly to investigations of mathematics and science instruction (Guskey, 2003). Third, these results show that although the promise of research-based professional development remains largely unfulfilled, it need not remain so (Guskey, 2003)

Thus, according to Guskey (2003), there is no consensus in the professional literature about what constitutes the major criteria for defining effective professional
models. Further, there is little or no evidence of a professional development model that has empirically tested the hypothesis that the model positively impacts students’ growth in academic achievement.

Analysis of student learning data has typically shown a greater variation exists between classrooms within a school than between schools or between districts (Kifer, 2001). In other words, within the unique context of nearly every school, there are teachers who have found effective ways to help students learn. Identifying and finding ways to share the practices and strategies of these teachers among their colleagues might provide a basis for highly effective professional development within that context (Guskey, 2003). These data suggest that a critical feature of any professional development delivery model is the development of a collaborative learning environment at the building level for participating teachers (Guskey, 2003).

Despite the growing body of literature that supports the relationships among staff development, teaching quality, and student learning, some educators and policy makers question the value of providing time and resources for professional learning (Killion, 2002). However, many educators, including principals and teachers, embrace the link between student achievement and teaching quality and advocate for improved staff development. Elmore (in Farrace, 2002, p. 40) perhaps said it best: “If you’re going to make the changes in student learning that accountability requires, you have to dramatically increase the skill and knowledge of teachers and principals.”

**Students’ Achievement Growth**

Despite the amount of literature on professional development, relatively little systematic research has explicitly compared the effects of different forms of professional
development on teaching and the resulting impact on student learning (Desimone, Porter, Garet, Yoon, & Birman, 2002). Furthermore, most studies of professional development have not examined its effects in a quantitative and replicable manner (Desimone, Porter, Garet, Yoon, & Birman, 2002). Showing empirically that professional development translates into gains in student achievement poses tremendous challenges e.g. the nature of student achievement targeted in some professional development is not easily measured by most standardized tests, time is an issue for gathering real evidence of change, and the lack of longitudinal research are a few of the challenges (Richardson and Placier, 2001; Supovitz, 2001). This is despite an intuitive and logical connection between professional development and student achievement (Borko, 2004; Loucks-Horsley & Matsumoto, 1999; Supovitz, 2001).

Very few professional development programs provide any data as proof of their efficacy in improving student or teacher outcomes even as standards for evaluating staff development programs are being developed and refined (Guskey, 2002; Killion 2002). Accountability is the defining headline for education today (Jerald, 2000; Lewis & Shaha, 2003). In the past, organizations could spend funds on programs that looked good and felt great. Cleverness and creativity may have been adequate criteria for program selection, and satisfaction of participating teachers may have been the main measure of success. Improvements in student performance or achievement were arguably considered implicit. Today, in the age of accountability, expenditures must represent investments that promise tangible improvements in teacher and student performance as verified statistically through data (Belzer, 2003; Guskey, 2002; Parry, 1996; Todnem & Warner, 1993).
Authors of the article, “Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement,” reviewed more than 1,300 studies identified as potentially addressing the effect of teacher professional development on student achievement in three key content areas (reading, mathematics, and science) (Yoon, Duncan, Lee, Scarloss, & Shapley, 2007). Only nine met the What Works Clearinghouse evidence standards, attesting to the scarcity of rigorous studies that directly examine this link. In a similar study, Davis and Thompson (2005) in their review of articles and reports on evaluations of technology-related professional development located more than 100 documents. They “eliminated the majority of the work because the evaluation described focused upon the lowest level of Guskey’s categories or did not use carefully constructed/tested instruments or failed to report clear results” (Davis and Thompson, 2005, p. 4). The fact that these two reports eliminated so many articles suggests the strong need for a clear evaluation design, the development and use of valid and reliable instruments, and the use of multiple data sources (Guskey 2000, pp 8-10). These reviews “clarify the need for creating a model that allows for the collection of scientifically based evidence and provides guidance for how to do this” (Davis & Thompson, 2005, p.2).

**Teachers’ Growth**

Authors of a recent paper note that in the United States there are no national data that examines professional development over time or links professional development participation to both changes in teaching practice over time and student achievement (Olson, Desimone, Le Floch, & Birman, 2002). These researchers suggest that to evaluate and improve our policies supporting teachers’ professional growth, we need to understand how professional development translates into changes in teaching practice and improved student
achievement (Olson, Desimone, Le Floch, & Birman, 2002, ). The review of this literature indicates that one reason for the limited evidence of the impact of professional development on student learning is the lack of longitudinal research on teacher change and student learning (Guskey, 2000; Cohen & Hill, 2000; Richardson & Placier, 2001; Supovitz, 2001; Garet et al., 2001; Guskey & Sparks, 2002).

This review of the literature points out that despite the lack of agreement on a common set of characteristics of effective professional development and a lack of empirical data, some research results indicate and most educators believe that organized and effective professional development does have an impact on student learning (Guskey, 2003). In order to implement professional development more effectively, the Iowa Department of Education (2002) developed an approach that it felt had the best chance of successfully promoting educational reform and thus improve student achievement in Iowa. This approach developed in 2002 is known as the Iowa Professional Development Model (IPDM).

**Iowa Professional Development Model (IPDM)**

The IPDM, serving as the system of delivery for professional development in Iowa, is embedded within the state’s other educational reform efforts including the Comprehensive School Improvement Plan and the District Career Development Plan. As can be noted in Figure 2, while student learning defines the effectiveness of professional development, there is also a defined cycle of professional development. As is typical in academic models, the cycle reflects activities engaged in during an academic year. The cycle of professional development consists of an initial planning phase that reflects curriculum development. The second phase is an ongoing phase that includes the three components of (a) staff training and learning opportunities, (b) collaboration and implementation of curriculum elements, and (c)
formative student/teacher data collection. The third phase involving the collection of summative data on both teachers and students, and these data serve as the basis for program evaluation efforts (Iowa Department of Education, 2002).

Figure 2. Iowa Professional Development Model

Technology in Support of Professional Development
The IPDM supported by technology would provide educators with an opportunity to interact with each other both within and outside of their school building, thus ending their isolation. By helping teachers to integrate their pedagogical practices and share with others much of what they learn through professional development, the probability increases that their professional development will lead to growth in student achievement (Birman, Desimone, Porter & Garet, 2000; Elmore, 2002). The primary means of achieving this goal is through the building of a professional learning community of grade/subject/similar school teams connected by technology that provides coaching and feedback (Joyce & Showers, 1995). In these communities, teachers would engage in discussions and observations (Sykes, 1999). These technological communication tools would increase the sustainability of professional development throughout the academic year. This model contrasts with the typical practice of professional development workshops that are single episode events with no follow up or support. To demonstrate that technology-supported professional development is both efficient and effective, one could develop a web-based teacher data collection system that would provide quantifiable data for the development of a structural equation model (SEM).

A review of published reports examining educational finance and performance by a Kentucky legislature subcommittee, found that most of these studies define efficiency as the maximum performance for any given level of resources. The general definition of effectiveness is the ability to achieve stated education goals. (Seiler, Ewalt, Jones, Landy, Olds, & Young, 2006) For this study, Efficiency is defined as a faster, more cost effective means of delivering and sustaining professional development activities. Effectiveness is
defined as providing the basis for determining the impact of professional development activities and student achievement in mathematics.

Summary

Today’s classroom environment is in transition as a result of the development of newly defined standards for both curriculum content and student performance. This requires a massive effort of intense professional development for teachers. There is a considerable body of professional literature on the effectiveness of professional development. Much of it represents widely agreed upon expert opinion, or hypotheses, about the effectiveness of particular programs and their link to student learning. Unfortunately, there is little empirical evidence of a causal relationship between professional development and gains in student achievement. This study attempts to develop a research design based on the IPDM supported by technology, to provide empirical evidence that professional development does impact student learning.
CHAPTER 3. METHODOLOGY

Purpose of the Study

The purpose of this research study was to develop a field experimentation model that would provide policy makers and educational professionals in Iowa schools with a method (or approach) for evaluating the impact of professional development on student achievement. Specifically, the study assessed the impact of technology-supported professional development delivered by an Area Education Agency’s (AEA) professional development team to middle school mathematics teachers in central Iowa who then implement instructional strategies in order to promote student growth in mathematics achievement during the eighth grade. Further, the study attempts to provide empirical evidence that technology-supported professional development activities improve student learning.

Research Hypotheses

The current study identifies professional development as the educational intervention (independent variable) that is introduced to impact (cause) improvement in student mathematics achievement (dependent variable). The student data collection design was a pretest/posttest design with a randomly selected comparison group. The unit of analysis is at the building level. The teacher data collection design was a time series design with teachers reporting the implementation frequency of curriculum components (pedagogical strategies) on a monthly basis during the academic year. The research questions and research hypothesis for the study are as follows:

A. Do professional development activities conducted within the context of the Iowa Professional Development Model (IPDM) improve student achievement?
**Research Hypothesis 1.** A longitudinal cohort of students in schools where teachers receive professional development during the 2006-2007 academic year will demonstrate greater growth in mathematics achievement than students in schools randomly selected where teachers did not receive the treatment.

**Research Hypothesis 2.** A longitudinal cohort of students in schools where teachers received professional development in mathematics during the 2006-2007 academic year, will demonstrate greater growth in their mathematics achievement than in their reading achievement.

B. Does the use of technology to create professional learning communities support teachers’ implementation of professional development activities?

**Research Hypothesis 3.** A Structural Equation Model\(^3\) (SEM) can be developed that demonstrates the impact of various elements of the cycle of professional development on student achievement gains during the academic year.

To place the research hypotheses in proper context, the design logic must first be considered. Hypothesis I is related to research question 1? Do professional development activities conducted within the context of the Iowa Professional Development Model (IPDM) improve student achievement? Answering the question involves a comparison of eighth grade teachers in schools receiving mathematics professional development (the experimental

---

\(^3\) Structural equation modeling is a statistical methodology that takes a confirmatory (i.e. hypothesis-testing) approach to the analysis of a structural theory bearing on some phenomenon. Typically, this theory represents “causal” processes that generate observation of multiple variables. The term structural equation modeling conveys two important aspects of the procedure: (a) that the casual process under study are represented by a series of structural (i.e. regression) equations, and(b) that the structural relations can be modeled pictorially to enable a clearer conceptualization of the theory under study. The hypothesized model can then be tested statistically in a simultaneous analysis of the entire system of variables to determine the extent to which it is consistent with the data. If goodness –of-fit is adequate, the model argues for the plausibility of postulated relations among variables: if it is inadequate, the tenability of such relations is rejected. (Byrne, 2006, p.3)
group) with eighth grade teachers in schools not receiving mathematics professional development (the control group). Data collection and analysis for the experiment took the form of a pretest/posttest design where the pretest preceded the educational intervention and the posttest followed the educational intervention. If the gain in student performance for the experimental schools receiving the educational intervention were greater than the growth in mathematical achievement of the comparison group, the impact of mathematics IPDM is demonstrated. For this study, the data collected and analyzed were eighth grade students’ scores on the ITBS administered in the spring of the seventh grade and the spring of the eighth grade. The educational intervention activities, implemented during the 2006-2007 academic year, occurred between the pretest and the posttest.

The second research hypothesis is an extension of the first research hypothesis: “does a longitudinal cohort of students in schools where teachers received professional development in mathematics during the 2006-2007 academic year, demonstrate greater growth in their mathematics achievement than in their reading achievement (as measured on the ITBS)? To test this hypothesis, data were collected and analyzed only from students in the experimental group of schools; i.e., those students who received the educational intervention and were used to compare growth in mathematics with growth in reading achievement. In this data analysis, each student serves as his/her own control.

Having addressed questions pertaining to the overall effectiveness of the IPDM, we turn to the third research question: “does the use of technology to create professional learning communities support teachers’ implementation of professional development activities?” A Structural Equation Model (SEM) can be developed that demonstrates the impact of various elements of the cycle of professional development on student achievement gains during the
academic year. In this SEM analysis both student and teacher data were analyzed in an attempt to develop a causal model focusing on the impact of IPDM elements on gains in student mathematical achievement during the eighth grade. The intervention elements analyzed were: (a) teacher use of technology for sustaining building learning teams, (b) teacher implementation of educational intervention activities, and (c) teacher perception of school support. Teacher data were collected via a web-based instrument. Items from this instrument are shown in Appendix A.

**Research Design**

The experimental group of schools consisted of 11 self-selected middle schools located in Heartland Area Education Agency (AEA 11), one of ten regional educational service units for the state of Iowa. During the 2006-2007 academic year, these schools were in their first, second, or third year of participation in the E$^2$ T$^2$ project. Consequently, while students were experiencing the eighth grade mathematics intervention strategies for the first time during the 2006-2007 academic year, their teachers had varied experience with implementing the strategies. A total of 1,377 students in the experimental schools participated in the study.

The comparison group was made up eighth grade students in 30 randomly selected schools from around the state of Iowa that had been and statistically weighted to be representative of the state defined school size categories; e.g., Category 1 is composed of school districts with fewer than 250 students, and Category 7 is composed of the state’s

---

4 E$^2$ T$^2$ project was the State of Iowa’s name for activities conducted under the Title II D or Enhancing Education Through Technology of the NCLB. The funding from this program was used to finance the activities conducted by the various applicants. These activities were focused around elementary reading and mathematics and middle school reading and mathematics.
largest schools with enrollment of 7,500+ students. A total of 2,339 students in the comparison group participated in the study.

Fifteen teachers in the 11 aforementioned schools participated in the data collection effort by reporting frequency of strategy implementation, frequency of technology and learning community activities, and perceived degree of school support. Teacher data were used only in the causal model analysis.

**Achievement Measures**

Student achievement data served as the dependent variable for the growth curve analyses and student achievement gain scores served as the criterion variable in the causal model used in this study. For the between-group growth curve analyses comparing experimental and comparison schools, the mathematics total score served as the dependent variable. In the within-subject growth curve analysis, each student’s mathematics achievement gain was compared with his/her reading achievement gain during the eighth grade. Thus, the mathematics total score and reading total score served as the dependent variable for this analysis. In the third analysis, the causal model, student gain scores in mathematics achievement during the eighth grade served as the dependent or criterion variable.

The National Standard Score (NSS), is the product of the Hieronymous\(^5\) scaling technique employed during the development of standardization norms for ITBS, is a value that describes a student’s location on an achievement continuum (Kolen, 2006). This standardized score is the basic score from which all other performance scales associated with

\(^5\) Hieronymous scaling, also called grade-equivalent scaling is the procedure used to develop the score scale for the Iowa Tests of Basic skills.
the ITBS (i.e., Iowa percentile rank, national percentile rank, and the grade equivalent score) are derived. The range of NSS scores associated with typical eighth grade performance is 200 to 300.

**Educational Intervention**

The overall goal of Heartland’s educational intervention activity was that all students would become proficient in mathematics. With this in mind, there were only three variables that could change: the student, the curriculum, and/or the teacher’s behavior. Since the students came as who they are and curriculum decisions are made at the district level, the focus of Heartland’s efforts was on changing teacher behavior to improve student achievement.

Based on their review of the research on effective mathematic instruction in middle schools, Heartland believed a quality mathematics curriculum has five interwoven strands that accurately reflect the complexity and dynamic nature of learning mathematics with deep understanding. These five interwoven curriculum strands are: (a) conceptual understanding, (b) productive disposition, (c) procedural or computational fluency, (d) adaptive reasoning, and (c) strategic competence (National Research Council, 2001, p. 5).

Heartland’s project was grounded in the implementation of five research-based instructional strategies incorporating the previously mentioned five curricular strands at the middle school level. Based on a task analysis of the mathematics curriculum at the eighth level, the five research-based instructional strategies that promote the five identified curriculum strands are: (a) establishing routines (mental mathematics and daily mathematics review), (b) providing worthwhile tasks, (c) encouraging student discourse, (d) extending student thinking, and (e) basing instructional decisions on student understanding (National
Research Council, 2001, p. 5). The focus of this study was to examine the impact of teacher implementation of the strategies and the use of technology as a tool for communication and support of professional development.

**Data Collection Design**

Student data were collected from a cohort of students for whom there were ITBS composite scores for reading and mathematics for both 2005-2006 and 2006-2007 school years. The 2005-2006 test scores represented the pretest. The 2006-2007 test scores were used as the posttest. Scores for students were used if the students attended schools that met the following criteria: (a) schools (both experimental and comparison group schools) that tested during the spring semester, (b) schools that continued with the same testing semester (e.g., spring testing in 2005-2006 school year and spring testing again in 2006-2007 school year), and (c) experimental group schools that required teachers to complete the E²T² teacher survey during the 2006-2007 school year.

Teacher data collection relied on teacher reporting to a web site. Teachers logged onto the web site with their ID–folder number and their individualized password. Teachers were expected to report on their activities once per month. The teachers reported on their implementation of intervention strategies, professional learning communities (PLC) activities, the degree of support they received from building administration or AEA staff, and amount of technology use. Teacher data are nested under buildings and used only in the causal model analysis.

**Summary**

The study was designed to collect data that would help answer the research questions linking professional development activities to student achievement. The design included data
collected from teachers and students from an experimental group as well as student data from a comparison group. The purpose was to measure the impact of systemic professional development activities and the implementation of the teaching strategies that were enhanced through the use of technology on student achievement.
CHAPTER 4. ANALYSIS OF DATA

Professional development as planned and delivered to schools by the Area Education Agency through the IPDM must be implemented in the classroom in order for student achievement gains to be observed. This study’s evaluation effort was situated in eighth grade mathematics. Since not all of the elements of the IPDM could be evaluated in a single study, this study focused on the system components of (a) professional development, (b) teacher collaboration/implementation, and (c) student growth in mathematics achievement during the eighth grade. Analyses of data related to research question one and its two hypotheses were conducted using SAS software. Analysis of data related to research question two and its hypothesis was conducted using Mplus 5.1 software. The following results are presented in the order of the research questions addressed.

Research Question 1

Research question 1 was, “Do professional development activities conducted within the context of the Iowa Professional Development Model (IPDM) improve student achievement?” Connected to this question were two hypotheses:

Research Hypothesis I

A longitudinal cohort of students in schools where teachers receive professional development during the 2006-2007 academic year will demonstrate greater growth in mathematics achievement than students in schools randomly selected where teachers did not receive the treatment.

Research Hypothesis II
A longitudinal cohort of students in schools where teachers received professional
development in mathematics during the 2006-2007 academic year, will demonstrate greater
growth in their mathematics achievement than in their reading achievement.

The initial analysis, a test of Hypothesis 1 employed a mixed design 2 by 2 Analysis of Variance (ANOVA) with repeated measures. The first factor is a between-group factor with 2 levels (i.e., the experimental and comparison groups). The second factor is a within-subject factor with 2 repeated measures (i.e., the pretest and posttest of students reading and mathematics scores). The test of hypothesis I rests with an inspection of the interaction of the two groups. Specifically, greater gain in mathematics achievement for the experimental group would produce a pretest-posttest growth curve with a greater slope than that of the growth curve for the comparison group. This would be reflected in a statistically significant two-way interaction. The null hypothesis would predict no difference in slope (i.e. a non-significant interaction). Student achievement data were total mathematics scores based on student scores from the mathematics concepts and estimation test, the mathematics problem solving, and data interpretation tests from the eighth grade form of the ITBS.

The rationale for using a mixed design ANOVA with repeated measures to analyze differences in rate of improvement in mathematics achievement during the eighth grade of the 2006-2007 academic year is as follows. When analyzing data collected from years other than the first year of an educational intervention, the researcher must take into account the fact that the effects of whole building professional development are cumulative for both teachers and students. For purposes of this study, whole building professional development for middle school mathematics teachers involved sixth, seventh, and eighth grade mathematics teachers.
The curriculum activities that served as the educational intervention for this study were geared to eighth grade mathematics content. Data collected for the current study reflected the second year of professional development for teachers from some buildings and the first year of professional development for teachers in other buildings. Thus, it could be expected that a main effect for level of achievement at the beginning of the eighth grade would exist for the experimental schools. However, this effect difference was not a problem given the mixed design ANOVA with repeated measures because of the co-variance procedure involved in the data analysis. Thus, the rate of improvement was not biased in favor of the experimental group even though the two groups started at significantly different levels of curriculum implementation.

**Hypothesis I**

For this analysis, the experimental group consisted of 11 schools and 1,377 students. The comparison group consisted of 30 schools and 2,336 students. Results indicated the main effect for the year was statistically significant $F(1,39) = 134.71$, $p<.01$; as was the main effect for the group $F(1,39) = 9.69$, $p<.01$. However, the interaction was non-significant, $F < 1.0$.

As noted in Figure 3, students in both the experimental group and the comparison group improved at approximately the same rate during the eighth grade. The mean NSS scores for the experimental group increased from 265.11 (sd = 36.02) to 274.91 (sd = 37.23). This finding represents a gain of 9.8 NSS units. The mean NSS scores for the comparison group increased from 251.07 (sd = 37.83) to 261.47 (sd = 38.71) which represents a gain of 10.4 NSS units.

Although the predicted significant interaction was not observed, the gain by the experimental schools can also be view from an added value perspective. The experimental
group demonstrated an average Grade Equivalent gain of one year and two months (GE = 1.2) of growth in mathematics achievement during the eighth grade.

**Figure 3. A Comparison Pretest Posttest NSS scores of Experimental and Comparison Schools**

![Graph showing comparison of NSS scores between experimental and comparison schools](image)

The significant main effect for group reflects a mean of 270.01 (sd = 36.63) for the experimental group and 256.27 (sd = 38.27) for the comparison group when collapsing across pretest and posttest. This higher level of mathematics achievement at both pretest and posttest is likely attributable to the fact that some participating teachers were in the second or third year of professional development activities focusing on mathematics achievement. In this respect at the beginning of the year (pretest) the mean national percentile rank was 74 for the experimental group of schools and 62 for the comparison group of schools.

The significant main effect for year reflects a significant gain in mathematics achievement for both groups during the eighth grade. When collapsing across groups, the improvement from pretest (M = 258.09, sd = 36.93) to posttest (M = 268.19, sd = 37.97)
indicates a significant improvement for schools regardless of group (experimental or comparison).

**Hypothesis II**

To test hypothesis 2, a two by two within-subject ANOVA was used to compare growth rates in mathematics achievement and reading achievement for students during the eighth grade. This analysis involved only the 1,377 students from the 11 experimental schools. For this analysis, each student served as his/her own control; this analysis is viewed as a finer grained analysis. It is the two-way interaction that is of interest. Again it would be assumed that since mathematics was being targeted by the curriculum interventions and reading was not, that more rapid growth should be observed in mathematics achievement during the eighth grade.

Results were similar to the findings of tests for hypothesis 1. The main effect for the year was significant $F(1, 4119) = 587, p< .01$. As noted in Figure 4, when collapsing across groups, the posttest mean of 269.24 (sd = 33.15) was significantly higher than the pretest mean of 258.16 (sd = 31.98). Of special interest is the subject main effect $F(1, 4119) = 757.9, p< .01$ (Table 3). Collapsing across pretest/posttest, mathematics achievement was significantly higher (Mean = 269.99, sd = 36.65) than reading achievement (Mean = 257.44, sd = 28.49). Thus, by the second or third year of the mathematics intervention initiative, students’ level of mathematics achievement was significantly higher than their reading achievement.

The rate of gain in mathematics and reading achievement was tested in the two-way interaction. The year by subject interaction was statistically significant $F(1, 4119) = 8.16, p< .01$. However, in this case, although reading achievement was at a lower level through the
eighth grade, the average rate of improvement was greater for reading (NSS gain score = 12.35) than for mathematics (NSS gain score = 9.76). Means for the two-way interaction are reported in Table 1. Although the gain in mathematics achievement was occurring at a less rapid rate than reading achievement, in terms of grade equivalent scores, students gained 1.2 years in mathematics achievement during the eighth grade.

**Figure 4. Mathematics and Reading Scores for Experimental School Students**

![Graph showing mathematics and reading scores comparison between 2005-2006 and 2006-2007](image)

**Table 1. Means and Standard Deviations for the Two-way Year by Subject Interaction**

<table>
<thead>
<tr>
<th>Subject</th>
<th>2005-2006</th>
<th>2006-2007</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Mathematics</td>
<td>265.11</td>
<td>36.04</td>
</tr>
<tr>
<td>Reading</td>
<td>251.26</td>
<td>27.93</td>
</tr>
</tbody>
</table>
Table 2. Within-Subject ANOVA Results

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>1</td>
<td>167908.00</td>
<td>167908.00</td>
<td>587.42</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Subject</td>
<td>1</td>
<td>216642.00</td>
<td>216642.00</td>
<td>757.91</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Year*Subject</td>
<td>1</td>
<td>2331.95</td>
<td>2331.95</td>
<td>8.16</td>
<td>&lt;.0043</td>
</tr>
<tr>
<td>Building</td>
<td>10</td>
<td>154113.00</td>
<td>15411.00</td>
<td>4.58</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Student (building)</td>
<td>1,363</td>
<td>4586718.00</td>
<td>3365.16</td>
<td>11.77</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Residual</td>
<td>4,119</td>
<td>1177383.00</td>
<td>285.84</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Research Question 2

Does the use of technology to create professional learning communities support teachers’ implementation of professional development activities?

Hypothesis III

Research Hypothesis 3. A Structural Equation Model (SEM) can be developed that demonstrates the impact of various elements of the cycle of professional development on student achievement gains during the academic year. Hypothesis 3 addressed the question of how selected elements of the Iowa Professional Development Model (IPDM) worked in promoting student gains in mathematics achievement during the eighth grade. Data, collected from participating teachers, were analyzed using a structural equation model (SEM) with a path analysis approach. The analysis was not an attempt to demonstrate that the educational interventions delivered by the IPDM caused the change in student mathematics achievement during the eighth grade. This issue was addressed when testing hypotheses 1 and 2. Rather, the SEM approach was used to determine the impact of selected elements of the IPDM on student achievement gains in mathematics during the eighth grade. IPDM elements tested were: (a) activities supported by technology, (b) building principal support, (c) AEA support
of school staff, and (d) teacher learning community activities. The teacher reporting form is detailed in Appendix A.

Results reflect what occurred in the 11 experimental schools. Means, standard deviations, and zero-order correlations for the five measured variables are shown in Table 3. The multivariate normality test was administered to determine whether the data met the normality assumptions underlying the maximum likelihood of the procedure used to test the models in the present study. Results of the multivariate normality test developed by Mardia (see Bollen, 1989) indicated that the data were not multivariate normal, \( \chi^2 (2, N = 1377) = 1.32, p > .05. \)

Therefore, the scaled chi-square statistic developed by Saorra and Bentler (1988) was used for adjusting the impact of non-normality on the results.

The maximum likelihood method in the Mplus 5.2 program (Muthén & Muthén, 2007) was used to conduct the path analysis. As suggested by Hu and Bentler (1999), three fit indices were used to assess goodness of fit of the models: (a) the comparative fit index (CFI) whereby values of .95 or greater indicate that the model provides an adequate fit to the data; (b) the root-mean-square error of approximation (RMSEA) whereby values of .06 or less indicate an adequate fit; and (c) the standardized root-mean-square residual (SRMR) whereby values of .08 or less indicate an adequate fit. As noted above, the scaled chi-square was used to adjust for the impact of non-normality on the results (Satorra & Bentler, 1988).

---

6 Multivariate normality is required by maximum likelihood estimation (MLE), which is the dominant method in SEM for estimating structure (path) coefficients. Specifically, MLE requires normally distributed endogenous variables. The Bollen-Stine bootstrap and Satorra-Bentler adjusted chi-square (this is an adjustment to chi-square which penalizes chi-square for the amount of kurtosis in the data. That is, it is an adjusted chi-square statistic, which attempts to correct for the bias introduced when data are markedly non-normal in distribution. These are used for inference of exact structural fit when there is reason to think there is lack of multivariate normality or other distributional misspecification. See also Bollen (1989).
The corrected scaled chi-square difference test (Satorra & Bentler, 2001) was used for the comparison of nested models.

Initially, a fully recursive model was tested (see Figure 5). Two paths (from learning community 0607 to post-math 8 and school support 0607 to post-math 8) were not significant ($\beta = -.01$ and .01, $ps > .05$). This finding indicates that neither learning community nor school support directly impacted student gains in mathematics achievement. However, the impact of these two elements of the IPDM was mediated by teacher implementation. Consequently, these two paths were trimmed.
Figure 5: The Fully Recursive Model for Heartland Data

Mplus 5.2 was used to analyze this model. All path coefficients are standardized scores. Goodness of Fit Statistics are acceptable: \( \chi^2 (0, N=1377) = .00; \) CFI= 1.00; RMSEA = 0.00; SRMR= 0.00.

**p < .01; *p < .05
The trimmed model was tested, and results are shown in Figure 6. The model fit was acceptable, $\chi^2 (2, N = 1377) = 1.32$, $p > .05$; CFI = 1.00, RMSEA = .00, SRMR = .00. As expected, results supported the hypothesis that both professional learning communities and school support positively impact strategy implementation in mathematics teaching ($\beta$s = .25 and .18, $p < .01$). This finding indicates that the more frequently teachers used professional communication, and received support during the school year, the more effectively they implemented intervention strategies in their classes. Although, strategy implementation as a mediator variable impacted gains in mathematics achievement, the path did not reach statistical significance ($\beta = .03$, $p > .05$). Descriptive data for model elements are reported in Table 4.
Figure 6. Trimmed Model for Heartland Math Experiment 8th Grade (Midyear and Spring) N=1,374

Mplus 5.2 was used to analyze this model. All path coefficients are standardized scores. Goodness of Fit Statistics are acceptable: \( \Delta \) scaled \( \chi^2 \) (2, \( N=1377 \)) = 1.32 (\( p > .05 \)); CFI = 1.00; RMSEA = .00; SRMR = .00.

\*\*\( p < .01 \); \*\( p < .05 \)
Table 3. Means, Standard Deviations, and Intercorrelations for Measured Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Pre-Math06</td>
<td>265.11</td>
<td>36.02</td>
<td>---</td>
<td>-.07**</td>
<td>.01</td>
<td>.03</td>
<td>.78**</td>
</tr>
<tr>
<td>2. Learning Community07 (q3)</td>
<td>.29</td>
<td>.45</td>
<td>---</td>
<td>.02</td>
<td>.25**</td>
<td>-.06*</td>
<td></td>
</tr>
<tr>
<td>3. School Support07 (q4)</td>
<td>.88</td>
<td>.33</td>
<td>---</td>
<td>.18**</td>
<td>.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Strategy Implementation07</td>
<td>2.09</td>
<td>1.22</td>
<td>---</td>
<td></td>
<td>.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Post-Math07</td>
<td>274.91</td>
<td>37.23</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes. N = 1,377. Pre-Math06 and Post-Math07 = 0506 and 0607 Math NSS total scores from Iowa Tests of Basic Skills (ITBS). Q3 denotes Question 3 on the teacher questionnaire: Has your school established a Professional Learning Community around this E²T² initiate? Q4 denotes Question 4 on the teacher questionnaire: This month, have you received encouragement and support for the E²T² initiative from your building administrator? **p < .01; *p < .05.

Thus, results of the SEM provides a basis for further study of the effectiveness of the model in providing educational interventions that facilitate educational reform of the Iowa public educational system.

Summary

The analysis of the data for the first research question did not provide any significant difference between the student growth for the comparison and the experimental group. Both groups grew at about the same rate. However, the trimmed model’s (see Figure 6) finding indicates that the more frequently teachers used professional communication, and received support during the school year, the more effectively they implemented intervention strategies in their classes. While not significant it does provide a model for using data to support the impact of professional development on student learning.
CHAPTER 5. DISCUSSION

Empirical Data Model

Discussion of Research Question 1 and Hypothesis I

Did professional development involving the training and monitoring of teachers’ implementation of educational interventions have an impact on student achievement in mathematics? As pointed out in the review of literature, there is a general lack of empirical data demonstrating that professional development impacts student achievement (Davis & Thompson, 2005; Yoon, et al., 2007). The focus of the current study was the development of a quasi-experimental data collection model that would address this question. The proposed model used growth curves to measure the impact of the educational interventions. This was accomplished by using longitudinal cohort data in a two-group pretest-posttest design whereby the educational interventions (in this case, professional development training and implementation) were treated as the independent variables and the pretest and posttest mathematics scores as the dependent variables. The two groups were a treatment group and a comparison group. The impact of professional development on student achievement during the eighth grade was measured relative to level of achievement and rate of achievement.

In terms of level of achievement, a significantly higher level of achievement at both the beginning and the end of the eighth grade (e.g., significant between-group main effect) was observed. Thus, eighth graders taught by teachers who had participated in the professional development activities for a year or two performed better than students whose teachers had not received or implemented the professional development activities.
Although students in the intervention schools achieved at a higher level during the eighth grade, the rate of growth in mathematics achievement was no more rapid for them than for students in the comparison schools. In other words, there seemed to be an accumulative effect of whole school professional development activities from the seventh grade to the eighth grade; i.e., students who experienced the implementation of the intervention strategies with seventh grade materials appeared to start the eighth grade at a higher level and maintained that advantage throughout the year. Thus, the claim can be made that the implementation of the professional development activities by participating classroom teachers was a possible cause of the observed difference in level of mathematics achievement during the eighth grade.

**Discussion of Research Question 1 and Hypothesis II**

A second way of addressing the growth in mathematics achievement during the eighth grade was to treat each student as his or her own control and compare level and rate of achievement in mathematics and reading achievement performance for each student. While this is one way of drawing a comparison, it begs the question, “Does mathematics and reading achievement grow at the same rate”? Without a proper no-treatment control group, there is a confounding variable. Thus, this data collection design is weaker than the mixed design employed to test Hypothesis I.

Results of the within-subjects analysis confirm the cumulative effect of the professional development activities on level of mathematics achievement. As in the between-subjects main effect, the level of mathematics achievement was higher than the level of reading achievement at both pretest and posttest. Thus, the cumulative effect of whole school professional development training appears to be replicated from a different perspective.
Interestingly, there was a significant two-way interaction indicating that the rate of reading achievement during the eighth grade was more rapid than the rate of growth in mathematics achievement. This finding may be simply a regression to the mean effect due to the lower starting level for reading achievement at the beginning of the eighth grade. It could also be the result of a reading intervention that may have been conducted at the same time as the mathematics intervention. Regardless, due to the aforementioned confounding variable, this should be viewed at best, a weak replication of the rate of growth question.

However, from an added value perspective, the rate of growth in mathematics achievement during the eighth grade by students in schools receiving professional development was quite positive. When the gain was translated into a grade equivalent score, the average gain in mathematics achievement for nine months of instruction was one year and two months of growth.

*How the Model Works - Discussion of Research Question 2 and Hypothesis III*

Having examined the impact of teacher implementation of the professional development strategies and the rate of mathematics achievement during the eighth grade, the SEM model was used to describe how links are forged between the implementation of professional development training and student achievement.

The SEM model was designed to reflect activities engaged in during the academic year by participating teachers from the experimental schools. Activities being monitored were (a) frequency of technology use in the development and sustainability of a mathematics learning community of teachers, (b) teacher support from the school principal, and (c)
teacher implementation of strategies. The path analysis provided an opportunity to determine if professional development training activities directly impacted student achievement gains.

The fully recursive causal model indicated that neither technology/learning community activities nor principal’s support directly impacted student achievement gains. In other words, using technology, building learning communities, and having a supportive building principal did not directly improve student achievement in mathematics. Rather, these building characteristics directly impacted the implementation of the strategies by the classroom teacher. As can be noted from the trimmed model, the greater the use of technology to support learning communities the greater the teacher implementation throughout the school year. The same relationship was found in terms of principal support. Interestingly, the impact of technology/learning community development was greater than the impact of school support defined as a supportive principal. Thus it can be said with conviction that a supportive teaching environment contributes directly to implementation, and the strategy implementation can be viewed as the mediating variable for student achievement gains observed when testing Hypotheses I and II.

Next, the nature of the relationship between teacher implementation and the observed gains in student achievement in mathematics was examined. Results indicated that the greater the frequency of implementation, the better the student gain in mathematics achievement. However, the level of impact did not reach statistical significance, suggesting that the impact between strategy implementation and student achievement gains could well be an artifact of (a) how the implementation data were collected and prepared for statistical analysis, and (b) the fact that students and teachers are nested within buildings.
For purposes of statistical analysis, implementation data for each teacher were averaged across the intervention strategies. With this approach, findings for some interventions might indicate a positive impact while findings for other interventions might indicate no impact or even a deleterious impact. Thus, this averaging of the teacher survey data could be viewed as a possible statistical artifact. Future research efforts could provide for the analysis of single strategy impact on student achievement. This may require a re-conceptualization of the SEM model as a two-stage model whereby the first stage would examine the impact of building characteristics on teacher implementation and the second stage would examine the impact of teacher implementation on student gain. This approach would provide an overall impact of the educational intervention (as was done in the current study) as well as the impact of individual strategies.

A two-stage SEM model would have potential benefit for stakeholders such as AEAs that are responsible for the development of educational interventions and professional development training. Results would provide AEA teams with information that would enable them to evaluate the effectiveness of the prescribed educational interventions.

The second issue, the nesting of students and teachers within buildings, may also contribute to low impact between teacher implementation and student achievement. This nesting effect essentially means that each building has a single teacher implementation value (grand mean). The solution to this problem is relatively simple. Collecting and analyzing classroom data including student and teacher names could address the nesting issue. Politically, this solution would also require a culture change whereby such data are readily available.
Conclusions

Two conclusions can be drawn from this study. First, using the proposed model, general gains in 8th grade mathematics achievement are promising. Gains were consistently at or above the natural growth that was to be expected. Second, while there was no significant difference in the gain scores for students, the results indicate that the IPDM as a professional development delivery system – with the elements of theory, demonstration, practice, coaching and feedback – is a promising system for changing teacher behavior (Joyce & Showers, 1995) and impacting student achievement.

The IPDM as used in this study allows teachers and principals to engage in professional learning communities involving teachers from other schools and content experts from the professional development provider who interacted both physically and via technology with team members. These communities, composed of subject and grade compatible teachers, engaged in discussions varying from face to face meetings to discussion boards and IP video conferencing. These discussions via technology seem to have moved the project teachers along the levels of transfer from imitative use to integrated and executive control (Joyce & Showers, 2002).

Shortcomings of the Study

This does not mean there were no problems during the study. Technical problems such as use of out-dated browsers, limited bandwidth at the school, lack of IT (Information Technology) support, and/or IT network control and filters that blocked content and teacher interaction may have diminished the overall effectiveness of the technology support. Data collection was another problem. While data were collected only once a month, some teachers felt there were so many demands on their time that they had difficulty finding time to fill out
the eight-question survey. Lapses in reporting data may also have been the result of teachers’ or schools’ fear of accountability or comparison. In addition, teachers may have wanted to provide only socially desirable responses to survey questions. Therefore, their self-reported responses may not have accurately described their behaviors.

**Future Research.**

In the present study, there was no attempt to assess individual differences in teacher report of implementation. In future studies, teachers could be grouped into three levels of reporting frequency; e.g., high reporting teachers (i.e., those who responded consistently to the survey); and moderate and low implementers (those who reported implementations less frequently. It may well be that (a) consistency of reporting is a moderating variable that reflects differences in student achievement gains, and (b) consistency of reporting is a proxy for consistency of implementation.

The study also had problems with the structure of the teacher survey and the questions asked. The most of the questions asked were answered with either a yes or a no. This was done to make the survey as quick, simple and easy to complete for the teachers who felt pressed for time. This resulted in one question being used to determine learning community and one question being used to determine principal support. While this was an efficient and cost effective way of collecting information, it left a great deal to be desired. In future studies a more in depth survey could be attempted with participating teachers being rewarded somehow e.g., participating teachers being entered into a drawing for several technology prizes for their classrooms.

As a result of this study, it was noted that there is a need to validate the teacher survey data. To help correct this problem, it was determined that future studies should make
use of an electronic principal walk-through to validate teacher implementation and activities in the classroom. Principals would be using personal digital assistant (PDA) to gather data on any area; from instructional practices in the classroom to student engagement and lesson differentiation. The addition of these devices would create an additional data point in the model. Thus the model would have teacher survey data, electronically recorded principal observation data and student achievement data. This triangulation of data could then be used to; learn more about instruction and learning, validate effective practice and ensure continued use, create a community of learners for adults and children, focus teachers and the principal on student work and the learning process.

The study’s research design has potential for transforming education in the future. It allows schools to draw on the talents outside their immediate sphere and to collect empirical data linking professional development activities and student achievement. The research design demonstrates the potential positive impact that technology can have not only on professional learning communities but also on possible linkages between teachers out in the field, intermediate educational service agencies (AEAs), and policy developers (Department of Education). Institutions of higher education can also become a powerful partner with LEAs and intermediate agencies in supporting change in prek-12 education. For example, colleges could use this research design to create learning communities of recent graduates to support them, increase their contact with other new teachers, and break down their isolation as they start their teaching careers. In doing so institutions of higher education could collect data regarding the effectiveness of their pre-service program and perhaps help reduce the high teacher attrition rate among recent graduates. Ingersoll (2001) reported an average attrition rate of 46 percent among new teachers after six years. The design of the project
could also increase higher education’s understanding of the professional development activities conducted by the LEAs and intermediate education agencies and thereby create a basis of discussion regarding their current teacher preparation program.
APPENDIX

Teacher Survey

1. In your lessons, are you implementing the instructional strategies or activities that you received in your E2T2 professional development sessions? YES  NO

2. Is computer or IP technology being used to support your implementation of these E2T2 strategies and activities? YES  NO

3. Either in the building or via technology, has your school established a Professional Learning Community around the E2T2 initiative? YES  NO

4. This Month, have you received encouragement and support for the E2T2 initiative from your building administrator? YES  NO

5. This month, have you received encouragement and support from your Area Education Agency E2T2 professional development team? YES  NO

6. On average, how frequently do you implement the **Base Instructional Decision on Student Understanding/Base Instructional Decision on Student Understanding** strategy in your classroom? [Select] (Never, once per week, 2 times per week, 3 times per week, 4 times per week, 5 times per week (Daily))

7. How much time do you spend (in minutes) implementing the **Base Instructional Decision on Student Understanding/Base Instructional Decision on Student Understanding** specific instructional strategy during a typical class period? [Select] (None, 1-10 minutes, 11-20 minutes, 21-30 minutes, 31-40 minutes, 41-50 minutes, 51-60 minutes)

The last two questions were repeated for each of the six middle school strategies.
REFERENCES


*Implementation research: A synthesis of the literature* (FMHI Publication No. 231). 

Tampa: University of South Florida, Louis de la Parte Florida Mental Health Institute, 
The National Implementation Research Network.


Does one size fit all? In P. G. Andrews & V. Anfara, Jr. (Eds.), Leaders for a 

movement: Professional preparation and development of middle level teachers and 

administrators (pp.145-160). Greenwich, CT: Information Age Publishing.


Teachers College Press.


Leadership, 61*(7), 42-46.


*Review of Educational Research, 47*(1), 335-397.


ED354588)

Furhman, S. H. (Ed.). (2001). *From the capital to the classroom: Standards-based reform in 
the states*. Chicago: University of Chicago Press.


professional development effective? Results from a national sample of teachers. 


