Cleveland State University
EngagedScholarship@CSU

ETD Archive

2010

An Empirical Investigation of the Role Of "Ownership" in Creating Sustained School Reform

Sharon A. Brown
Cleveland State University

Follow this and additional works at: https://engagedscholarship.csuohio.edu/etdarchive

Part of the Education Commons

How does access to this work benefit you? Let us know!

AN EMPIRICAL INVESTIGATION OF THE ROLE OF "OWNERSHIP"

IN CREATING SUSTAINED SCHOOL REFORM

SHARON A. BROWN

Bachelor of Science in Psychology

University of North Dakota

May, 1980

Master of Education in Adult Learning and Development

Cleveland State University

May, 1989

Submitted in partial fulfillment of requirements for the degree

DOCTOR OF PHILOSOPHY IN EDUCATION:

POLICY STUDIES

at

CLEVELAND STATE UNIVERSITY

DECEMBER, 2010

This dissertation has been approved for

the Office of Doctoral Studies,

College of Education and Human Services

and the College of Graduate Studies by

_____

R. D. Nordgren, Chairperson

     Curriculum and Foundations

_____

Joshua Bagakas, Methodologist

     Curriculum and Foundations

_____

Marius Boboc, Committee Member

     Curriculum and Foundations

_____

Paul Williams, Committee Member

     Counseling, Administration, Supervision and Adult Learning

_____

James Salzman, Committee Member
     Ohio University

ACKNOWLEDGEMENT

AN EMPIRICAL INVESTIGATION OF THE ROLE OF "OWNERSHIP"

IN CREATING SUSTAINED SCHOOL REFORM

SHARON A. BROWN

ABSTRACT

School reform policies have failed to produce sustained positive changes in education practice. Theories of school change provide structure to reform policy. Program evaluations focus on implementation and outcomes but seldom test the theoretical assumptions of the initiative. This study tested theory, specifically the influence of "ownership", against the experience of Reading First Ohio. This literacy program was implemented in elementary schools in low performing and very low socio-economic status urban and rural public districts in Ohio from 2003-2009. Archival data were analyzed using structural equation modeling. The analysis confirmed that the constructs of leadership and classroom change are mutually critical elements in school reform. The model failed to identify specific variables within the initiative structure that were tightly aligned to the theoretical assumptions. This study has implications for strengthening school reform policy.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

In recent decades, school reform has become a politically charged national imperative prompted by agency reports (A Nation at Risk, 1983), social pressures (Berman & McLaughlin, 1979), and global competition (Friedman, 2005). With political interest came policy initiatives and the cyclic funding that follows. Federal policy has focused millions in financial resources on school change initiatives over the past several administrations. For example, in the year 2000, the Comprehensive School Reform Demonstration Program had an annual budget of $220 million dollars. Later, the No Child Left Behind Act (2001) invested $200 billion in education with a substantial portion focused on reform efforts (McGuinn, 2006). The large investments in school reform have prompted schools to embark on change efforts, ready or not. Farrell (2003) described government based funding as the primary catalyst for school change in many educational settings. He added that this process of cyclic efforts has created a revolving door of school reform efforts that align only with funding cycles. Fullan (2001) challenged funding driven reform and discussed innovation as the appropriate catalyst for change. Yet, school reform efforts and shortfalls cannot be dismissed as a cyclic funding or policy issues alone. Fullan (2001) appropriately described the current environment as

one where school reform innovations are abundant and schools have many good choices. Unfortunately, even in the event that schools recognize the need and are focused on the change they often do not have the organizational capacity to identify the best choices, implement with fidelity and sustain those efforts over time. Too often, schools begin an initiative, implement and when the funding cycle ends move on to the next. Lasting change, also referred to as continuation, institutionalization, and sustainability, has remained an elusive target even in the face of initially successful implementation efforts (Datnow & Stringfield, 2000).

Change in Schools

Diffusion theory provides a useful lens for examining the role of innovation in organizational change. Diffusion is "the process in which an innovation is communicated through certain channels over time among the members of a social system" (Rogers, 2003, p. 5). Diffusion of innovation requires that an organization recognizes a concern or problem, matches an innovation to the organizational agenda that addresses the issue, restructures the innovation to meet their needs, clarifies for further diffusion and finally makes the innovation routine in the life of the organization (Rogers, 2003). It is the ability of the organization to make the innovation a routine aspect of the life of the organization that makes the innovation sustainable. The theoretical steps appear simple: the real-world processes much less so. Educational change models can be either overly simple single-factor theories or they can be exceedingly complex, theoretic and unwieldy.

Even well-designed organizational change is fraught with challenges. Marris (1975) stated that all change involved anxiety, loss, and an inevitable struggle for the individuals involved. As a result, overt changes in process, actions, or behaviors are

resisted by both organizations and the individuals within them. According to Senge (1990), the tension that is created by change can either spur creativity and energy for the change process, or it can create a wave of resistance and "push back" against the proposed change. The degree to which an organization is able to manage the tension of change and direct it to a positive outcome dictates the overall success of a change process.

Educational organizations are inundated with reform initiatives that are launched and abandoned with regularity. Research has demonstrated that systems are not only resistant to initial change processes; they are highly resilient. Resiliency was displayed in the stubborn manner in which educational organizations returned to previous practices— often in far less time than was required by the change process itself (Datnow, 2005). Fullan (2001) described the many pressures applied in the initiation of an innovation process that often do not promote full implementation let alone long-term change. He stated that seldom are programs planned and implemented with deliberate intent to sustain and continue. Generally, the focus for schools is the implementation of model initiatives, not on long term or lasting impact.

Challenges to Studying Sustainable Change

Sustained school reform efforts are few and far between, and even when a program is sustained there is limited opportunity to learn from it. Too often programmatic efforts provide only for evaluation of the implementation of an initiative during and immediately after the program but not for long-term effects. Funding cycles, and the associated program evaluation, generally end with the pre-determined implementation cycle leaving few resources for the monitoring and data collection of a

long term sustainability study. In addition to the issues caused by short cycles of

implementation that align with funding, the literature discussed the issue of programmatic

fade. Programmatic fade occurs as a program slowly diminishes in implementation

quality and intensity over time due to lack of focus, resources, and/or staff turnover.

Programmatic fade undermines long-term implementation and certainly diminishes the

ability to study and learn about sustainability. Finally, studies of sustained school change

require long-term research efforts. Longitudinal studies have traditionally been limited

by time, resource, and subject availability and are less frequently conducted than other

research designs. (Bijleveld & Leo, 1998) As a result, sustainable school reform is an

area of concern for many but researched by few.

Reading First as a School Change Model

Reading First is the early literacy component of the No Child Left Behind Act

(NCLB). NCLB authorized over $900 million in funds for K-3 reading improvement

focused on practices based in scientifically based-reading research (SBRR). Reading

First Ohio is the Ohio implementation of the federal Reading First K-3 literacy program.

The goals of Reading First Ohio:

1. To establish a high quality primary classroom reading programs anchored
   in scientifically based reading research;

2. To help teachers acquire the knowledge and skills they need for effective
   instruction, sound instructional decision-making, accurate diagnoses and
   powerful interventions that ensure children's progress;

3. To coordinate resources in the service of reading success for all children
   (USDERF, 2002).

The program provided specific compliance guidelines from the federal grant but allowed significant autonomy of implementation for districts in the state of Ohio, a local control state. The federal program mandated specific classroom time commitments to reading (90 minutes daily), professional development (180 minutes monthly), specific staffing configurations (Literacy specialist, Data manager, Resource coordinator, District coordinator), SBRR appropriate core reading program (selected by the district), Dynamic Indicators of Basic Early Literacy Skills (DIBELS) assessment on specific benchmarks time schedule and end-of-year assessment using Ohio Achievement Test (OAT) in Grade 3 and Terra Nova in Grades K-2. The required compliance efforts created a baseline of implementation expectations for all districts involved in the initiative. Fidelity of implementation at the compliance level was monitored in multiple ways through the grant.

Reading First as a reform effort focused on implementation and capacity building at the district, building, and classroom levels. The key programmatic factors were substantially held constant across all sites. Each site was expected to sign Ohio Department of Education assurances confirming their intent to implement with fidelity. Although key program factors were held constant, evolution and customization of the program was promoted in Ohio through the process of data-based technical assistance, program development, materials selection and professional development. Core program materials were reviewed by district leadership teams and selected to meet the needs of their district. Professional development was customized to the needs of the individual districts within the parameters of the program focus. Teachers were instructed in the use

of student data to differentiate instruction and to provide intervention for students within the classroom and through specialized interventions.

Characteristics of a Quality School Change Initiative

Diffusion theory discussed the processes of moving an innovation into and through an organization. Fullan (2001) summarized the literature in the area of characteristics of high quality innovations for school reform. He stated that an innovation needs to have four characteristics; organizational need, clarity, complexity and quality. Rosenblum and Louis (1979) identified the degree to which the school system had formal recognition of unmet needs as a key factor of readiness in subsequent implementation of innovation. The alignment of recognized school needs and the new innovation allows school personnel to perceive that needs are being addressed thus strengthening the likelihood of full implementation and sustained effort (Bodilly, 1998). Districts identified as eligible for Reading First participation were underperforming in the area of reading achievement and are all high poverty districts. The low reading performance of the students in these districts was a significant concern to the Ohio Department of Education as well as to the district administration and teaching staff. Reading ability is a widely accepted fundamental for all students (National Reading Panel, 2000) making Reading First more easily accepted by the individual buildings and districts. Also, districts were required to secure acceptance by the teachers' union prior to application for the initiative, which was designed to reinforce that the need had been identified and shared by personnel beyond central office.

Clarity in the innovation is another characteristic of a quality change model. Unclear goals, undefined expectations, and non-specific steps for implementation

undermine the possibility of successful change (Fullan, 1999). Gross et al. (1971) found that in a study of a large school reform effort the majority of the teachers who were presumed to be involved could not explain the basic features of the innovation. In addition to substantial financial resources, Reading First Ohio provided clear guidelines, definitions and implementation expectations for districts and buildings. Implementation was monitored on a regular basis through the Program Monitoring Tool (PMT) and technical assistance was provided to insure that clarity of purpose was maintained. Classroom teachers were provided with ongoing professional development and classroom coaching that reinforced the common language and expectations of Reading First Ohio.

According to Fullan (2001), complexity of the innovation refers to the extent of change required within the organization or asked of the individuals involved as well as the overall difficulty in implementing the innovation. Unfortunately, the literature conflicts in this area. Highly complex innovations were found to make greater demands on the individuals implementing and therefore risked lower rates of success (Berman & McLaughlin, 1977). However, complex innovations were also found to have a greater impact and, therefore, deeper meaning for the participants which increased the likelihood of commitment (Fullan, 2000). Reading First Ohio is a complex innovation requiring districts and individual elementary buildings to engage in multiple levels of professional development, data collection and analysis, as well as changes in instructional practices but the importance of the effort and clarity of purpose facilitated most districts in reducing implementation challenges. Also, because Reading First focused on an area of instruction that was already central to the early elementary curriculum it reduced some of the difficulty of implementation.

Quality and practicality of the program as well as the quality of previous implementation efforts was found to be indicative of the successful implementation of reform efforts. In short, organizational capacity for change is built through experience and that capacity strengthens future efforts but capacity building takes time. Fullan (2000) discussed the political reality of adoption and implementation of innovations. He stated that quick adoption and short timelines for implementation are a common concern for reform efforts. He discussed the challenges of overcoming previous shortfalls or negative attitudes toward change caused by low quality innovations, impractical approaches or compressed timelines. As a complex innovation, Reading First Ohio had a relatively short implementation timeline. To overcome some of the challenges of the compressed timeline, support structures were built into the grant in the form of specific personnel roles at the district level, technical assistance at the state level and professional development support at the state level. Reading First Ohio required application for participation in late winter and early spring with the implementation initiating the following fall. While the timeline was short, the substantial technical assistance and support provided to participants ensured that the implementation of the model was not compromised due to the timeframe. All teachers in grades K-3 were provided with ample professional development to insure their understanding of both the grant and scientifically based reading research. Program monitoring was conducted early in the implementation, which allowed a clear picture of the implementation and more immediate intervention for concerns. In addition coaching, technical assistance and support were provided directly to building personnel during startup to minimize frustration on behalf of personnel participating in the implementation.

As a model for school reform Reading First Ohio met the basic characteristics of a quality school reform model as described in the reform literature. Despite efforts to provide technical assistance, clarify the model, and provide clear guidelines for implementation, districts participating in Reading First Ohio achieved at dramatically different levels. Program monitoring provided evidence of common basic levels of implementation, student data provided evidence that achievement improved, and yet there were discernable differences in the depth of program implementation. It was apparent that those differences would impact the likelihood of sustainability.

Theoretical Framework for School Change

According to the literature, the factors that most impact sustainable reform efforts include:

1. Quality model or initiative, as described previously (Fullan, 2001).

2. Classroom change brought about through teacher professional development in the areas of instruction and curriculum, teacher professional collaboration, demonstrated student success (Fullan, 2000, 2001; Newman &Wehledge 1995).

3. Leadership and leadership development (Barker 2006; Nettles, 2007; Robinson 2007; Wetherill 2005).

These three core factors provide the basis for the first level of a sustainability model. In Reading First Ohio, these three areas were the basis of improved literacy in Grades K-3. Reading First Ohio focused significant time, professional development, technical assistance, and resources in assuring that the basic model of the Reading First Ohio

initiative was diffused to all teachers in Grades K-3, and all administrators in the selected buildings.

Unfortunately, Reading First Ohio occurred during the same years as a major new school construction and reconstruction project in the State of Ohio. While the excitement of new buildings was positive it also created some challenges. Districts were on a schedule established by the state and had no ability to impact the timing of the construction. As a result many Reading First buildings and districts experienced significant barriers to implementation due to construction such as moving students, redistricting, building changes and consolidation of buildings. Additional challenges emerged in the form of teacher and leader turnover over the course of the grant. As part of the sustainability model for this study, challenges were defined as turnover, building construction and redistricting that resulted in shifts in student and personnel location disrupting relationships that had been built, and multiple layers of bureaucratic structure that impeded communication and decision making.

Beyond the first level implementation factors that create the components of this change model are the interactive processes that move an innovation toward ownership: Adaptive Collaboration as a variable and an organizational action, presumes a full understanding and implementation of the model or innovation and then the active adaptation of that implementation. This aligns with Rogers (2003) phase identified as *adaptation for additional diffusion* and began the process of making innovation routine in the organizational structure. In the case of Reading First, this is the point where school sites were actively engaged in on-going, vigorous, data-based decision making about teaching, instruction and intervention that extended the depth and intensity of the work.

Ownership is both a variable and in this model, a proxy for sustainability. Ownership of both the model and the processes within the organization are identified by efforts to expand beyond the boundaries of the chosen model while retaining the identified priority components of the model. The expansion occurs through the efforts of the building stakeholders and in response to their professional growth in the implementation process as well as the identified needs of the organization. In RFO this is identified as moving the RFO model beyond Grade 3 or extending the programmatic processes into additional content areas.

Ultimately a sustainable change effort is signaled by evidence of ownership of the change model in the form of an extension or adaptation of that model despite any barriers that may be introduced to the model. Fullan (2001) stated that: "True ownership is not something that occurs magically at the beginning, but rather is something that comes out the other end of a successful change process" (p. 92).
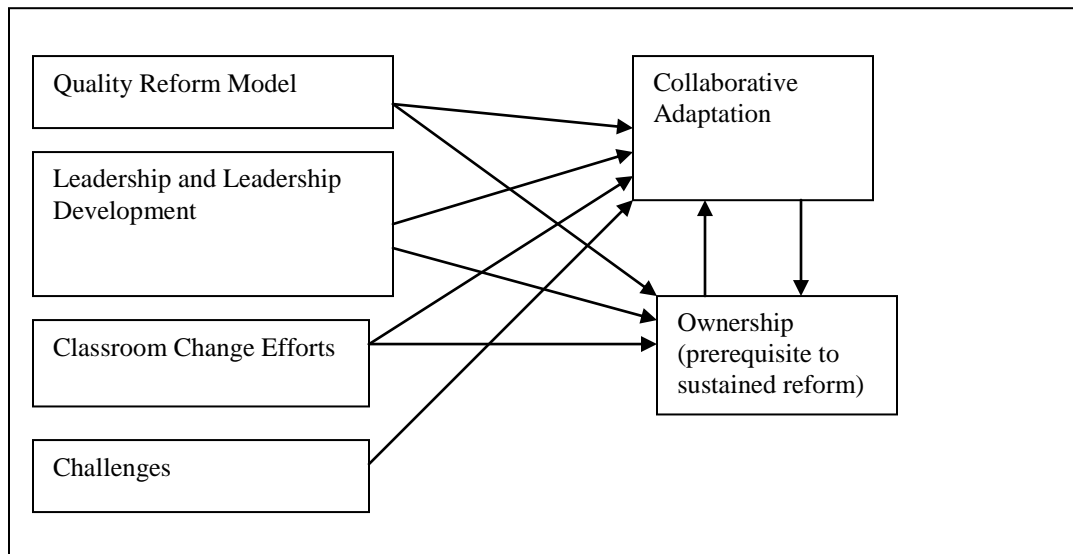
Figure 1.Basic theoretic framework of proposed sustainable school change

Statement of the Problem

Education invests millions of dollars in school reform innovations, yet the results of change efforts remain inconsistent and unpredictable. Even reform efforts that are fully implemented have little likelihood of being sustained. There is scant literature that focuses on prerequisite behaviors and programmatic actions for sustainable school change. There is a need to understand the programmatic path to sustained change so that future policy initiatives can build in the necessary supports for long-term, positive change. The purpose of this investigation is to test a theory-based model of sustained school change and the contributory constructs against the empirical evidence of Reading First Ohio.

General Research Questions

1. To what extent does the theory-based sustainability model accurately reflect the reform experience of the identified Reading First Ohio schools?

2. To what extent does the model accurately predict the districts most likely to sustain their reform effort as identified by the degree of ownership demonstrated?

3. To what extent does is the model able to accurately predict the performance of the schools relative to expert judgment?

4. To what extent do the latent variables of ownership and collaborative adaptation impact the overall functioning of the model?

5. To what degree do the identified barriers obstruct the ability of the school/district to achieve a fully implemented and sustainable model?

Significance of the Study

This study is significant in its contribution to understanding the issue of sustainable school change and the body of literature surrounding it. The assumption of all school reform is that once change is achieved, schools will never go back (Fullan, 2000). Unfortunately, education is rife with examples of programs successfully implemented and quickly discarded. Too often successful outcomes and measures of program implementation are considered the summative product of a program implementation. Successful outcomes and full program implementation are necessary but not sufficient indicators of lasting change (Datnow, 2005). In this investigation, program implementation and successful outcomes will be incorporated as variables in the larger issue of sustainable change efforts. Reading First Ohio data, years 2003-2009, will be used to build a theory-based model of sustainable school change demonstrating that a prerequisite to sustainable efforts is the ownership of the change process by school personnel and to test the hypothesized model against the progress of Reading First Ohio schools toward sustained efforts. The purpose of this study is to build on the body of literature on sustainable school change by providing empirical evidence of the contributory nature of ownership.

Delimitations

This investigation is limited in several ways:

1. The study is based on Reading First Ohio data available from years 2003-2009 only.

2. Ownership as a construct is serving as both a prerequisite to and a proxy for sustained school reform due to the multiple funding programmatic.

extensions and accountability for local education agencies (LEA) from the

Ohio Department of Education.

3.   All definitions of implementation, change, and sustained efforts have been

operationally defined through Reading First and Reading First Ohio.

Operational Definitions

District type. As defined by Ohio Department of Education, there are three types

of districts that are included in Reading First:

1= Rural: High poverty

4= Urban: High poverty

5= Major Urban: Large student population, very high poverty

Ownership. For the purposes of this study, ownership is defined as the process of

adapting and/or extending the given school reform model (RFO) to grades or content that

are clearly beyond the scope of the original grant.

Professional development. Learning experiences to improve knowledge, skills,

and abilities of teachers, principals, and specialized roles within Reading First Ohio.

Professional development is the primary vehicle for change in classroom practice and

leadership development in Reading First Ohio. Classroom coaching is included as

professional development.

Program monitoring. Program Monitoring is a process of expert review of site

evidence in response to and guided by the Program Monitoring Tool (see Chapter III for

details). Review occurred three times per academic year and was the primary vehicle for

shaping technical assistance at the district level. Programmatic technical assistance is

included as part of the program monitoring process.

Reading First Ohio (RFO). Reading First is a federally funded program in early literacy. In Ohio this program targeted the financially poorest districts that had the lowest achievement scores on the state assessment of reading in Grade 4 (later Grade 3). The specific model for the basic program was designed and disseminated by the state as part of a grant process. Local implementation was the responsibility of the LEA.

School reform. Any school change process designed to improve school culture, student performance, instructional practices, curricular design, leadership, and/or any combination of the above through a specified plan of action, model or theoretic structure. Generally, also includes a specified timeframe.

Student achievement. Student achievement is demonstrated by student scores on the DIBELS, Terra Nova, and the Ohio Achievement Tests (see Chapter III for more detail.)

Sustained school reform. Sustained reform is a school change effort that extends beyond the cycle of anticipated implementation.

Technical assistance. A feedback and professional learning processes conducted through the collaborative review of programmatic data by expert providers and site-based Reading First personnel teams.

Summary

Global economic changes have spurred political pressures on education systems. Education reform models have multiplied in quantity and accelerated in the anticipated speed of implementation. Individual theorists have reviewed and discussed specific components of school reform implementation such as the role of leadership, of teachers, of students and of the model itself but our understanding of sustainability remains

limited. Using the experiences of Reading First Ohio as the base programmatic model

and data source, this study incorporates the research on organizational change to

empirically test a model of school reform sustainability focused on the role of ownership

as a prerequisite construct for sustained change efforts.

CHAPTER II

REVIEW OF THE LITERATURE

The following is a review of the literature focusing on organizational change theories and school reform. School reform literature is a collection and synthesis of reform history, policy and implementations as well as programmatic evaluations and outcomes. Study of reform history demonstrates shifting emphasis of the reform efforts, increases in the rate of reform cycles and the volume of reform initiatives. Further, review of policies and the implementation expectations demonstrates a shift from emphasis on inputs in the form of curriculum and instruction to a focus on implementation processes and finally to accountability for outputs. Much of what we know about school reform emerges from programmatic research and evaluations that focus on specified programs and implementation of those programs in a specific context. School reform remains a challenging issue for schools and researchers alike.

Early Change Efforts in Educational Organizations

For the first hundred years of public schooling in America, change in schools consisted primarily of gradual shifts in programs, processes or services designed to accommodate changes in society or population. Lasting changes in education tended to be long in coming and incremental in nature (Tyack, 1995). For example, prior to 1870

kindergarten was reserved for the few, but as population grew, concerns over early childhood preparation shifted. The demand grew until, by the 1960s, kindergarten became universal (O'Neil, 2000) and now, in 2009, universal, public-funded preschool is being proposed (Stephans, 2009). In 1900, only 6% of teens attended high school; by 1960, 70% of teens earned a high school diploma (Ravitch, 2000). High school graduation was once considered unnecessary for the general population (Tyack, 1994), now the current president is calling for postsecondary schooling for all (Obama, 2009). These examples evolved and "institutionalized" over the course of a century.

Early change efforts were primarily local in nature and gradual in pace. In contrast, current reform efforts are often state or federally driven, bounded by funding cycles with substantive changes expected in a short time. While policy shifts in education have prompted change efforts from the 1960s onward, researchers are bereft of evidence that demonstrates a clear path to lasting positive change (DeBray, 2006).

Decades of School Reform

The 1950s were a relatively stable decade in education, with innovations focusing primarily on instructional practices and expanding educational attainment for the majority population (Fullan, 2001). In 1957, *Sputnik* was launched and the space race contributed to public interest in universal education. In addition, the *Brown vs. Board of Education* decision of 1954 and the civil rights movement of the 1960s moved American education from two systems, one black and one white, to one system expected to educate all children. Throughout the 1960s and 1970s access and equity remained focal reform issues as schools struggled to adapt to the changes in society. Academics disputed best pedagogy, best curriculum, and best practices in education. Urban districts became the

focal point for reform efforts as it became increasingly clear that poor and minority

children were not receiving an equitable education (Ravitch, 2000). In 1983, *A Nation at*

Risk prodded Americans to rethink public education and launched an era of accelerated,

largely externally driven and often rapidly shifting school change efforts with the

following challenge.

> Thus, we issue this call to all who care about America and its future: to
>
> parents and students; to teachers, administrators, and school board
>
> members; to colleges and industry; to union members and military leaders;
>
> to governors and State legislators; to the President; to members of
>
> Congress and other public officials; to members of learned and scientific
>
> societies; to the print and electronic media; to concerned citizens
>
> everywhere. America is at risk. We are confident that America can address
>
> this risk. If the tasks we set forth are initiated now and our
>
> recommendations are fully realized over the next several years, we can
>
> expect reform of our Nation's schools, colleges, and universities. This
>
> would also reverse the current declining trend--a trend that stems more
>
> from weakness of purpose, confusion of vision, underuse of talent, and
>
> lack of leadership, than from conditions beyond our control (NCEE,
>
> 1983).

A Nation at Risk launched a flurry of policy and programmatic responses at the

federal, state and local levels. "Through 1984, 1985, 1986 and beyond education was a

dominant issue in state capitals nationwide…In all there were an estimated 3,000 separate

school-reform measures enacted in the states during the mid 1980s" (McGuinn, 2006). In

this era reform ideas were rampant but without alignment or consistency. Successful schools were often encouraged to launch multiple programs to provide a "showcase" for other buildings. It was possible for a single school building to have multiple initiatives focusing on different content, different grade levels and even instructional styles within grade levels (Bender, Sebring, & Bryk, 2000). The term "Christmas tree school" was coined to describe a school with multiple shiny-bright-initiatives all over it. Christmas tree schools emerged as reform demonstration sites were encouraged to launch multiple programs and initiatives at the same time, unaligned and often with conflicting philosophies (Hatch, 2000).

A Nation at Risk called for upgraded academic standards for content areas as well as a specific set of "new basics" courses for high school graduation. A study by the U.S. Department of Education (1984) compared high school graduates from 1972 and 1980 only to find lower scores in both verbal performance and mathematics. The study also found that the students were doing less homework, grade inflation had increased and that the percentage of seniors taking an academic curriculum had declined 12%. "More and more students were going to college, even though they were not taking the courses necessary to prepare for college-level studies. By 1982, 50% of high school graduates went to college immediately after graduation, but only 9% of them had taken 4 years of English, 2 years of a foreign language, and 3 years each of social studies, science, and mathematics" (Ravitch, 2000, p. 410).

As a result of the pressure for increased academic standards for all students, the late 1980s and early 1990s focused on providing federal funds to states in order to build standards and accountability systems. States responded with the formation of committees

and stakeholder groups to study every aspect of the education system from content to length of the school day. Schools were expected to implement academic standards in addition to other changes identified on a state-by-state basis. An additional, and in many cases unexpected, outcome of the development of academic content standards and accountability systems was that it promoted reform efforts that could be documented and where the goals for student achievement could be standardized (Ravitch, 2000).

During the late 1980s and early 1990s, education became politically linked with economic competitiveness which served to keep school reform an urgent concern, politically charged and constantly in public awareness (McGuinn, 2006). American public education moved from a slowly evolving system of local changes and improvements to a national reform imperative fueled by federal and state interest in a host of narrowly focused and limited education change efforts. These efforts left schools engaged in multiple, often conflicting, efforts that stretched staff, resources and patience to their limit.

Whole School Change Efforts

Push-back on the multiple small and often disconnected reform efforts of the 1980s initiated a rise in whole school reform models. Large, prescriptive and externally driven reform efforts were launched in the form of the New American Schools project 1991, the Comprehensive School Reform program in 1998, Little Red School House in 1992, Successful for All in 1996, and others (Bodilly, 1998). The U.S. Department of Education's Strategic Plan for 2001-05 identified three major goals:

1.   To build a solid foundation of learning for all children.

2.   To reform the U.S. Education system.

3.   To ensure access for all to a high quality, postsecondary education, and life-long learning (USDE, 2000).

With reform as the second goal in the strategic plan, federal resources and attention focused intently on school change efforts.

The Comprehensive School Reform (CSR) initiative sought to demonstrate the best methods for producing school reform results through a series of model programs. CSR programs were expected to offer transformative change to schools throughout the country by providing model programs, creating model schools, and "scaling up" to other schools in the district, region or state. Models of school reform were identified and provided to schools as suggested approaches (Viadero, 1999). A proliferation of school reform models ensued. The Catalog of School Reform Models (Northwest Regional Educational Laboratory, 2001) provided a list of 63 improvement programs, 48 of which did not exist prior to 1980 – only one of which existed before 1960 (Hatch, 2001).

CSR was attractive to schools because, in addition to offering the prospect of improving education, the CSR models brought in external dollars (generally around $50,000 per building) and expertise to both support and monitor school change efforts (Keltner, 1998). Most major urban areas in the country launched multiple and varied whole-school reform efforts within their buildings. Overall, Comprehensive School Reform Demonstration (CSRD) funded 1800 schools nationally (Hatch, 2001) resulting in 380 school reform models initiated (Desimone, 2002).

The CSR models were viewed as advantageous over the smaller reform efforts because they eliminated the many piecemeal reform efforts and thus promoted a single focus within buildings, provided schools with external assistance and resources, and

finally, introduced clear models with quality control mechanisms (Keltner, 1998). The prescriptive nature of the whole school models created an opportunity for schools to both provide comprehensive professional development to the staff and to create a unified vision for improvement. CSR whole-school models were designed to be model programs that would be brought "to scale" and spread to other buildings. Unfortunately, most schools could not "scale up."Successful efforts were seldom replicated in other sites. Researchers were concerned by the inconsistency of the efforts and the tendency for schools to have multiple, unaligned initiatives that prevented a clear understanding of the issues surrounding replication or "scale up" (Bryk et al. 1998; Finn, 1997; Hill & Celio, 1998, Hatch, 2002).

Thus, the reform impact of CSR remained limited to single sites – often only one building or one grade level (McChesney, 2000; Stringfield, 2002). Multiple researchers raised questions about the sustainability of whole-school efforts (Bodilly, 1998; Datnow, 2005; Stringfield, 1998). Frustrated by the lack of clear results, researchers are left with mixed research results both on those efforts that did not work and for what was successful. Bodilly (1998) stated:

> Letting schools innovate on their own appeared to have limited success,
> resulting in the adoption of marginal programs, the disappearance of
> improvements when a principal or sponsor changed, or improvements in
> one or two schools but not many. Imposing state and district mandates
> appeared to offer similar meager success, with programs disappearing
> when state and district attention waned or when funding was reduced. The
> bottom line is that schools and districts have often faddishly adopted new

practices only to find them disappear within a short time or remain only in

a few selected schools in each district. Thus, a key frustration of those

who would improve schools has been the inability to translate the goal of

educating all students into coherent school level-responses within many

schools across the country or even within many schools across a district.

(pp. 1-2)

NCLB Era and New Reform Efforts

The No Child Left Behind Act (2001) maintained and expanded the focus on

education reform. NCLB added accountability to policy implementation. It modified

previous whole school reform efforts through a change in Title I that allowed funds to be

used school-wide. There were also expectations for student outcomes with academic

goals in reading and math. NCLB also launched content specific reform efforts in

reading (Reading First), math, and science (Math and Science Partnerships). NCLB

shifted the research and reform focus from school reform inputs in the form of whole-

school models to an emphasis on outputs in the form of accountability for student

achievement and school improvement. Media contributed to the interest in school

performance by regular publication of school accountability results (McGuinn, 2006).

Schools continued to launch one initiative after another. In a study of the Bay

Area District in California, 77% of the responding principals (over 51% response rate)

state that they had three or more improvement or whole-school reform efforts

simultaneously implementing in their building. Of these principals, 15% stated that their

building had six or more programs or initiatives. In a survey study of schools in both

California and Texas, 63% of the principals stated that their building had three or more

initiatives, 27% stated that there were six or more initiatives in their building. One

district in the study showed 18% of the respondents with nine or more initiatives (Hatch,

2002).

Reading First as a program of the NCLB was initiated in 2002 with an allocation

of approximately one billion dollars per year (USDE, 2002). Reading First was

implemented in 50 states, the District of Columbia, American Samoa, Virgin Islands and

The Bureau of Indian Affairs. The Reading First model provided structures that were

mandated of fund recipients while providing significant opportunities to adapt and

individualize the program. In Ohio, Reading First was implemented by a consortium of

three universities—Cleveland State University, The University of Akron, and John

Carroll University—called the Reading First Ohio Center (RFOC). The RFOC was

contracted by the State to deliver professional development and technical assistance to

local education agencies (LEA). The implementation of Reading First Ohio proceeded

from 2003 through 2010.

As the new millennia progresses, new waves of reform revolve around choice,

restructuring and the development of charter schools. Ohio has been in the lead of the

charter school movement, due in part to the leadership role of the Fordham Foundation.

A recent white paper provides the following policy recommendations as part of a report

on the state of things in Ohio: "Both McKinsey and Ohio Grant Makers Forum advocate

giving Ohio students and families access to high quality public school choices, both

inside and outside traditional systems." The study stated that nationally 30% of children

do not attend their neighborhood public schools and in Ohio's "Big 8" districts, attend a

school other than their local neighborhood school. School options include charters,

STEM schools, magnet schools, Early College Academies, and e-schools. E-schools, the most recent innovation, now attract 22,000 Ohio students (Fordham, 2009). The National Center for Education Statistics stated that in Grades 9-12, 50% of all courses will be taken on-line by the year 2019 (USCES, 2009).

The diversity of school reform choices and the rapid emergence of new forms of education delivery have served to accelerate availability and to vastly increase the consumer marketplace for education. Unfortunately, this has done little to insure quality in those choices.

> The genie of school choice is not going back inside the bottle, if only
> because family mobility and technology make it unstoppable. The
> emerging marketplace of schools in Ohio, however, is shadowed by huge
> variability in quality. The goal now must be to ensure that the quality of
> school choices keeps pace with their quantity (Fordham, 2009, pp. 17-18).

Although there are decades of school reform efforts past and future innovations pending, our understanding of the factors that create a quality and lasting school reform effort remain piecemeal.

Factors Contributing to School Change

Studies of past efforts have provided insights into some individual factors in educational change. Key roles within schools impact school reform efforts and have provided basis for professional development efforts for school personnel.

Role of building leadership.

Multiple researchers examined the role of leadership in the implementation and ultimate success of school reform efforts. Nettles (2007) identified the positive impacts

of quality school leadership on student achievement. This study demonstrated that leadership stabilized the culture of the building and overall contributed to improved instructional outcomes. Whitaker (2003) also focused on the relationships within the building, specifically the relationship of the leader with the teachers. He discussed the role of the leader as the one who builds a better teaching staff through new hires while also building the teaching staff that already exists. A study by PriceWaterhouseCooper (2007) discussed the problems when principals focus their energy on managerial tasks (either due to their own style or due to the system) instead of collaborative processes of teaching and learning. The most effective principals in terms of building culture and student outcomes were those who maintained a focus both administratively and through the building culture on teaching and student learning.

Even when the leader focuses on reforms that center on teaching and learning, the leader also risks becoming the sole change agent. Fullan (2001) discussed the leader's role in decision making in the change process. He cautions leaders about "false certainty" or over-confidence in programs that provide a prescriptive answer to the problems of the organization. Leaders can create their own "false certainty" within their own organizations if they become the change agent who is providing the cure-all. Fullan stated that this type of leadership prevents the organizational capacity building necessary to implement and sustain a successful reform effort. This type of change process becomes leader-driven and prevents a broader implementation buy-in and diffusion. One of the primary roles of a principal in a change initiative is to build an organization that is adaptive. Heifetz and Linsky (2002) stressed the importance of adaptive solutions. They discussed the limitations of technical solutions that focus on problems that can be solved

with current know-how. Adaptive problems require solutions that are not known. School leaders have to work with adaptive challenges on a daily basis; they have to build a team that can work to find adaptive solutions.

Robinson (2007) focused on the study of the interrelationship of the leader with the teachers. Robinson reviewed 17 studies with evidence of positive student outcomes and analyzed for descriptions of the leadership practices involved in each initiative. The analysis revealed five leadership dimensions that were critical in fostering teacher and student learning: providing educational directions; ensuring strategic alignment; creating a community that learns how to improve student success; engaging in constructive problem talk; and selecting and developing smart tools. The analysis showed that a distributive leadership approach to the improvement of learning and teaching is beneficial in a change process. Although distributive leadership was demonstrated as beneficial to change processes, it did not insure successful implementation or sustained efforts.

Finally, some leadership issues are basic. Research has shown that stable leadership was as important as the style of leadership. Barker (2006) studied the impact of leadership mobility on reform efforts and as expected, demonstrated that stable leadership plays a central role in reform efforts. Changing leaders diminished an organization's ability to move a reform process forward and to sustain organizational culture. A study by the Rand Corporation (1977) reviewed 293 federally funded projects and found that principal and staff turnover was one of the most significant factors in abandoning new reform efforts.

Role of central office leadership.

Central office is an often overlooked leadership component in school reform efforts. Whole school reform models focused on the building level actions and gave only limited attention to the role central office personnel played. However, studies have shown that without central office support reform efforts were unlikely to succeed. Rosenholtz (1989) studied leadership characteristics in "moving" districts as compared to those who were "stuck." Moving districts were those who had experienced some success in the implementation of a school reform agenda. By contrast, a "stuck" district was one that had also attempted to implement school change but had fallen short of a successful implementation. He found that superintendents of "stuck" districts were likely to blame principals for the failure of their building and to accept no responsibility for developing those same principals into instructional leaders. When districts take no responsibility for the development of principals, the principals become less likely to seek assistance to solve problems. Rosenholtz also stated that superintendents of "stuck" districts created a culture of self-reliance and professional isolation that made change efforts unlikely to succeed.

In addition to policies that do not support reform efforts, central office and state departments also contribute to or create barriers by policies that regularly shift reform priorities. Policies that chase funding and that encourage multiple change efforts in rapid succession create a cycle of reform efforts that cannot be fully implement or sustained at the building level (Little & Dorph, 1998, Spillane, 1999). Spillane (2000) further stated that even when the district endorsed and supported the reform efforts, lack of knowledge or understanding on their part often contributed to a weak, piecemeal and ineffective

program. He concluded that change efforts required capacity building at all organizational levels as part of the initial process of school reform. He stressed the need for central office personnel to be involved, informed and committed to any reform initiative.

Colburn (2008) discussed the role of central office in contributing to the ability of reform efforts to "scale up" by providing an alternative conception of "scale." She broadened the concept to include depth, spread, sustainability, and shift in reform ownership. Depth and spread are both processes that reflect the degree to which a staff has embraced and implemented a reform initiative. Sustainability of an initiative requires both leadership and policy shifts. She stated that inconsistency of policy and practice on the part of central office could undermine efforts to build capacity for sustained efforts. Ownership is the process of shifting from an externally driven program to one that is internally owned and directed. Colburn overtly stressed the need to link ownership and sustainability in school reform efforts and emphasized that one is unlikely without the other.

Role of teachers.

Fullan (2001) summarized the role of teachers simply; "all educational change depends on what teachers do and think" (p. 115). Changing what teachers do and think has been a challenge and the focus of years of professional development. Newman and Wehlage (1995) studied professional learning in the school reform efforts of Chicago schools. They found that investment in teacher professional development and collaborative efforts for teachers consistently netted positive impacts in the classroom. Working collaboratively on issues of teaching and lesson planning was proven to be an

effective method for improving student outcomes. Fullan (2000) further supported the value of professional learning communities; "If the district does not foster professional learning communities by design, it undermines them by default" (p. 165).

Professional learning communities (PLC) have been linked to student success by multiple researchers (Kruse & Louis, 1993, Newmann & Wehladge, 1995). However, it is important to note that although collaboration appears to be a vehicle to improved outcomes, Fullan and Hargreaves (1992) stated that it cannot be presumed that autonomy is negative and collaboration is positive. McLaughlin and Talbert (2001) found that while a weak professional learning community is always bad, a strong PLC can be either effective or not depending on whether or not they focus on getting results and improving practice.

In addition to building professional learning communities, it is important that those communities become data-driven. Hargreaves and Fullan (1998) discussed the need for teachers to become "assessment literate." They stressed the importance for teachers to learn to examine student data, to build capacity to act on the knowledge to improve instruction and school performance, and finally to effectively contribute to school efforts in reform.

Like leader turnover, teacher turnover negatively impacts school reform efforts. Darling-Hammond (1997) stated that a high percentage of teachers who leave the profession are those who have the characteristics that would make them excellent teachers. Thornton, Shipperson & Canavero (2007) found that teachers are most likely to leave districts and buildings where there are administrative issues and/or low levels of professionalism and professional community.

Attributes of the innovation or model.

Research also focused on the attributes of the innovations themselves that facilitate or impede success of reform efforts. Schools that adopted poorly designed efforts experienced disappointments in the implementation process (Timperley, 2005). The unique attributes of the innovation to be implemented play a role in the change process. Rogers (2003) stated that innovations must have particular attributes to be adopted. The innovation must be perceived to have advantages over other possible actions. It must be compatible with the ideals, culture, and experiences of the organization. Finally, the innovation must be perceived as having a reasonable level of complexity, able to be tried with limited risk and have some observable impact. Diffusion theory aligns with the school reform research. Fullan (2001) identified characteristics of the innovation to be implemented as critical to the success of the reform that are similar to the basics of organizational change literature. Characteristics such as a demonstrated need for the innovation, clarity in goals and processes of the innovation, appropriate levels of complexity to the innovation, and an understanding of the local and district contexts. Fullan provided a framework, but the application of these principles remains largely open to interpretation in the change process. Timperley and Parr (2005) added that common understanding of the innovation in the form of shared language and goals was necessary for any innovation to be successful.

Limitations of school reform research.

School reform research suffers from fragmentation both in content and in the focus of the research. The research can focus on multiple levels of reform; state (SEA), district (LEA), or building. The research can also focus on various roles within

organizations: principal, teacher, student, central office. There are researchers who focus

on specific reform models or specific contexts. There are research studies and there are

program evaluations. Hatch and White (2002) stated that the body of literature

surrounding school improvement is too fragmented to provide the basis for academic

discussion and clear understanding of school reform issues. "Despite widespread

conviction that adequate knowledge exists for improving schools, we argue that the

knowledge needed for successful school reform goes far beyond what is currently

available and accessible" (p. 117). In addition to the fragmentation, school reform

research suffers from several inherent limitations. First, school reform research is often

limited due to the potential bias introduced by the program developers researching their

own program. Due to the contextual constraints of school-based research, school reform

research seldom includes a control group (Pogrow, 2000; 2001). In addition, there are

challenges associated with researching across contexts of school reform. Finally, school

change is a dynamic process that is adapted as a process moves forward, which means

that strict research controls and definitions are often counter-productive and near

impossible (Cuban, 1993, Datnow & Stringfield, 2000; Fullan, 2000).

Early research efforts focused on understanding the implementation process itself.

Those schools that were quick to implement programs were studied to learn "what

worked" and those that were slow to implement were studied for "why not?" (Smith,

1997). The basic message of the research was to verify that buildings with strong

leaders, strong teacher communities, collaborative practices that focused on student

success and who selected a strong innovation were likely to implement with fidelity and

be successful in their efforts (Bodily, 1998; Datnow, 2000; Fullan, 2001; Smith et al. 1997; Stringfield & Ross, 1997; Stringfield et al., 1998).

In 2000, Datnow studied 16 projects, more than 300 case studies, where she demonstrated that reforms were more effective when educators at all levels (State, district and building) had common goals, and collaborated in the construction of action to reach those goals. Unlike much of the school reform research, this study examined key issues about reform implementation, independent of the change model being implemented. Datnow identified eight general factors necessary for high quality reform implementation: (a) finite set of shared goals, (b) goals that are tied to long-term, whole-team focus on school improvement, (c) a coordinated, broad-based plan for disseminating information about reform options, (d) schools engaged in thoughtful, critical process of inquiry about what needs to change in their school and why before they select reforms, (e)reform designs and reform designers must view local context and the diversity of the language, race, class and gender of those involved as strengths to build upon, (f)multidimensional, ongoing support, and leadership is required from design teams, district personnel and school site educators, (g)policy systems need to be aligned to support reform, and (h)successful implementation requires sensitivity and adaptability (without academic compromise) on the part of the design developers, local policy makers and educators in schools.

Individual features of the change process were also studied. Studies found that external school change coaches benefitted the reform process (Hall &Hord, 1987) but also could limit the capacity building within the organization. Too often in reform efforts, change agents along with school personnel were still learning the complexities of

reform and the roles each had to play in that reform effort (Hatch, 2002).   Hatch and his

associates ultimately highlight the need for building change capacity within schools.

While partner organizations can be beneficial in reform efforts, ultimately capacity must

be built within the organization if successful outcomes are to be sustained over time

(Hatch, 2001, 2006.)

Enthusiastic about the possibilities for education, researchers called for further

research into successful models of school reform. Unfortunately, the research failed to

provide clear insights and predictable paths for successful school change efforts.

Prerequisites for successful reform.

As an attempt to build capacity for school reform in districts and buildings,

multiple researchers developed lists of prerequisites to or attributes of successful school

change efforts to serve as guides for schools. Typically these lists consisted of sweeping

generalities about leadership, teaching, alignment, and support. Education Commission of

the States (1999) pulled together findings from implementing whole school reform. The

common learning from successful school reform stated that:

1. Comprehensive school reform changes the way schools, districts,

   and states do business (transform schools to focus on learning;

   district support is essential)

2. Legislative leadership sets the tone (provides a strong voice, builds

   coalitions, allocates resources)

3. State education department support is key for long-term success

   (schools in need are not necessarily school that can succeed on their

   own)

35

4.   Teachers make comprehensive reform possible(professional

        development is key; union support is vital)

5.   Evaluation - early and often – is critical (monitor, implementation as

        carefully as gains in student achievement; make results available to

        all)

Education Trust examined the student learning in 366 schools in 21 states and

found that the schools with the best student achievement shared six factors in their reform

implementation. The most successful schools used standards to design curriculum and

instruction. They devoted increased time to reading and math and provided professional

development focused on instruction. The schools initiated systems that monitored

student progress and provided intervention. They incorporated parent involvement

programs. Finally, all of the programs had accountability systems (Education Trust,

1999).

Hatch (2002) developed a list of attributes for the reform effort:

1.   If the school is not planning to stick with the reform, probably best not to

        start.

2.   Institutionalizing a reform requires hard work - requires time and

        sustained efforts.

3.   If principal and faculty believe that the reform is working well and value

        it, they need to keep district administrators informed about progress and

        successes so that they can advocate for the program, buffer it from change

        and criticism.

4.	Choice of reform efforts is important—one that supports improvement on

state and district measures of accountability is best.

The lists were too broad and too non-specific to provide clear insights into the change

process or a path to success. Districts and buildings were inundated with advice on

school reform but were still struggling to make it work.

Thus, although policies like the Comprehensive Reform Demonstration

Program look to improvement programs to help build the capacity to meet

higher standards, the implementation of these programs is difficult

precisely because the schools and districts in which they often work lack

the capacity to make changes and to coordinate varied initiatives (Hatch,

2002, p. 412).

Specific Challenges to Reform

While most researchers encountered and noted challenges to reform efforts, some

researchers focused exclusively on why initiatives failed. Mussoline (2001) found that

prescriptive whole-school models often did not take into account the contextual issues of

the school causing implementation problems and/or weak student outcomes. Failure to

acknowledge community, organizational climate and/or organizational culture were clear

barriers to successful school reform efforts and certainly to sustained efforts. Other

researchers focused on communication barriers. Implementation concerns were

attributed to a "disconnect" between central office and the building level implementation.

Buildings called for more support from central office while central office called for more

and explicit communication of needs by the building (Hatch, 2002; Spillane &

Thompson, 1997). Overemphasis on funding and funding cycles was identified as a

reform barrier (Fullan, 2001). Teacher and leader turnover were cited as concerns by multiple researchers (Hatch, 2002, Rand Corporation, 1977). Competing, conflicting, and opposing initiatives within the same district and/or building were cited by multiple researchers as a cause for implementation failure (Bryk et al. 1998; Finn, 1997; Hatch, 2001; Hess, 1999; Hill & Celio, 1998). Lack of clarity of the change innovation inhibited success. Researchers also found that the majority of the teachers in their study could not identify the essential components of the change innovation they were implementing. Inadequate resources and inadequate training (Darling-Hammond, 1994; Newmann & Wehlage, 1995) were found to negatively impact a school reform effort.

Ultimately, conflicting research produced questions about every facet of the change process (Hatch, 2002). The degree to which an initiative was able to address or avoid identified barriers was assumed to be a means to improve programmatic success, yet it did not insure success.

School Change Theories

Theories of school change emerged as researchers gained enough collective evidence to share perspectives on reform efforts and to address the concerns about low levels of impact or success in those efforts. Two major approaches to school change emerged; Concerns Based Approach and Process Focused Change Theory. One focuses on the person/change interaction the second focuses on the process/change interaction.

Concerns Based Adoption Model (CBAM).

Hall and Hord (1987) also discussed the concerns with incomplete implementation of innovations in educational organizations, not because of a weak innovation, but because the change-process itself was not sufficiently addressed. Their

theory, in contrast to model-specific CSR change theories, focused on the "personal side" of change. Hall and Hord focused on the interactions between the persons involved in change and the change process itself. They discussed the limits on innovation where model programs were left for the school personnel to implement as best possible. They proposed the Concerns Based Adoption Model (CBAM), which is primarily interested in addressing the personal and contextual issues that intercede in any change process. Early in their research, they described the phases of change as; research, development, diffusion, dissemination, adoption, implementation, institutionalization, refinement and abandonment. Over time, Hall and Hord refined CBAM and provided a more structured theory of design and implementation of change processes. Ultimately, Hall and Hord (2001) articulated their 12 principles of a school change process:

1. Change is a process not an event.

2. There are significant differences in what is entailed in development and implementation of an innovation.

3. An organization does not change until the individuals within it change.

4. Innovations come in different sizes.

5. Interventions are the actions and events that are key to the success of the change process.

6. There will be no change in the outcomes until new practices are implemented.

7. Administrator leadership is essential to long-term change success.

8. Mandates can work.

9.	The school is the primary unit for change.

10.	Facilitating change is a team effort.

11.	Appropriate interventions reduce resistance to change.

12.	The context of the school influences the process of change.

(pp. 4-14)

CBAM embraced the intricacies of change but in doing so, created a complex model that is difficult to translate into practical, actionable program processes.

Process Focused Change Theory.

Michael Fullan articulated a theory of organizational change specifically looking at school change. Fullan (2001) discussed school change focusing on the implementation of reform processes themselves. He described three major categories of the implementation process: characteristics of the change itself, local characteristics within the district and/or community, external factors such as government agencies. Like the previous theorists stated, Fullan emphasized that a focused, high quality and clearly necessary change or innovation will more likely be adopted. He also discussed challenges to implementation of a reform process. He found that there was a tendency to oversimplify the change process, which made prescriptive models attractive to schools and districts. Fullan discussed both the attraction to early comprehensive school reform models, and the reasons for failure. He described his rule of 25/75 wherein he stated that a good innovation is only 25% of the implementation issue, while a solid set of procedures and processes is the other 75%.

Research on Sustainable School Reform

Beyond implementation of reform efforts, researchers studied the possibility of long-term change. Datnow (2005) demonstrated that reform models that sustained were those that helped educators meet new local district and state demands, or at least did not conflict with them. In an earlier study (Datnow, 2000), she found that reform efforts that placed higher demands on the system and resources were less likely to reach sustainability. In a conflicting study, Fullan (2001) found that although reform efforts that were most comprehensive tend to be demanding on resources, they are also most meaningful to the school stakeholders and therefore most likely to be sustained.

Huberman and Miles (1984) stated that innovations are most likely to be sustained if they are institutionalized through policy and procedures. Rogers (2003) referred to this as a further diffusion of the innovation. Gladwell (2000) made the case that even the best change efforts, diffused through an organization could effectively be ignored if the factors necessary to create an ideological epidemic, or "tipping point," were not present. He stated that for any innovation to be sustained a critical mass of individuals must "buy-in" to the innovation and be committed to implementation. Datnow (2005) addressed the sustainability of comprehensive school reform (CSR) models in the face of turbulent district and state contexts. Her study drew on qualitative data gathered in a longitudinal case study of six CSR models implemented in 13 schools in one urban district. The author found that after 3 years, reform efforts ceased in nearly half of the 13 schools studied; two other schools were still implementing reforms but at very low levels. Only 5 of the 13 schools continued to implement their CSR models with moderate to high levels of intensity. The study demonstrated that changing district and state contexts affected the

sustainability of CSR models in schools differently depending on each school's strategy for dealing with the changes, such as their local conditions, experiences and reform and capacity. Others were dragged off course by demands from the school and community and where leaders were not effective in balancing demands of multiple stakeholders. High capacity schools with well institutionalized reform efforts did not experience the same sense of conflicting demands that the others did and therefore were not pulled off course. High capacity schools appear to have greater buy-in and ownership than those schools that were easily pulled off target.

Heath and Heath (2007) identified factors that facilitate the resiliency of change. The authors described the manner in which some ideas or changes achieve permanence that other, often stronger, ideas or concepts did not. According to the authors, simple ideas that are easily understood presented by a credible source with stories that had emotional appeal created fertile ground for continuity. Without "stickiness" ideas or innovations become a series of rotating changes that make little long-term impact. Whether it is called sustainability, ownership, or stickiness—lasting, positive change has remained elusive for education reform.

Ownership as precursor to sustainable reform.

Ownership is the product of group commitment and a focused change effort. (Coburn, 2003; Fullan, 2001). Coburn (2005) studied sustainability with a focus on scaling up. She concurred that buy-in for the innovation by those who will implement is a key to sustainability. She also discussed the related issue of a shift in ownership. Coburn stated that if an organization is to sustain innovations, there must be a critical

mass of buy-in accompanied by a shift of ownership from an external compliance to a personal ownership. Fullan (2001) extended his theory of program change by discussing the prerequisites for long term, lasting change. He detailed the supporting issues necessary for full implementation and ultimately a lasting change process. He stated that there must be active participation in the change process by school and district personnel, pressure to make and maintain changes, and support for the effort, authentic changes in behavior of the professionals – and finally, there must be ownership. "Ownership in the sense of clarity, skill, and commitment is a progressive process. "True ownership is not something that occurs magically at the beginning, but rather is something that comes out the other end of a successful change process" (p. 92).

Reading First Ohio as a Change Effort.

Reading First Ohio built upon the body of literature in several branches of education; reform, reading, leadership, instructional practices, and professional development. As a change model it incorporates many of the factors shown to be effective in the school change literature; professional development, leadership, clear guidelines and common language, accountability.

Reading First was focused on the National Reading Panel definition of scientifically based reading research (SBRR). The content of the professional development for teachers focused on the most current research in reading instruction. Grant requirements stated that teachers would receive 180 minutes of professional development monthly for the first 4 years of funding, and 120 minutes monthly in the final 2 years. Professional development was provided by literacy specialists and focused on reading instructional methods, data driven instruction, and interventions for struggling

students. The literacy coach served each building and provided in-class coaching, modeling of lessons, support for teachers and the site-based professional development (ODE Assurances, 2002; USDE, 2002)

Classroom implementation of Reading First required a 90 minute uninterrupted instructional block focused on reading. Reading instruction used SBRR materials that had been reviewed by and selected by district teams as part of the grant planning process. Students were assessed using the Dynamic Indicators of Basic Literacy (DIBELS) three times per year as a benchmarking process. DIBELS data were used to form dynamic groups and provided the basis for data-based instruction and intervention as needed. A combination of whole-group, small-group, and individual instruction was used within the classroom. End-of-year assessment of students consisted of Terra Nova in Grades K-2, and either the OAT or Terra Nova in Grade 3. OAT was not ready for use at the beginning of the grant (ODE Assurances, 2002).

Principals were an integral part of the programmatic process. Principals were provided with "Classroom Walkthrough Training" to facilitate their involvement with implementation of Reading First in the building. Principals also facilitated the building literacy team and served on the district leadership team (ODE Assurances 2002). District leadership was incorporated into the process through the role of the District Coordinator, a grant funded central office position that both administered the grant and served as an information liaison.

The factors evident in the Reading First Ohio site-by-site implementation serve as the independent variables within the model. The unique extensions and additions that

individual sites elected to incorporate into the Reading First Ohio model serve as the dependent variables in the proposed model.

Summary

The organizational change literature offers a framework for reviewing educational change in the larger context of organizations. Synthesizing the organizational change literature used to frame this review it is clear that common factors impact the change process regardless of the type of reform effort. The common factors in organizational and particularly school change:

- Clearly articulated, simple, and concrete goals facilitate the change process by providing common language and shared vision for the process.

- A leader who is able to align resources, personnel, and actions while also managing the politics of change is a critical element in the school change process. Collaboration between leader and teachers also contributes to the change process.

- Professional learning for teachers that identifies learning needs based on data is a fundamental component as well as teacher involvement in the planning and implementation processes.

- Accountability for positive student outcomes is a common research-based need for successful school change that will sustain and for implementation of the intended steps to reach the outcomes.

- Finally, shared communication of outcomes – or stories – in a manner that insures that stakeholders in the community, in student homes and in

45

political office feel the emotional and educational worth of the program is

necessary. This ultimately creates ownership of the change process.

This study will incorporate the factors identified in the literature into a statistical model

that will allow the empirical evidence of Reading First Ohio to be tested.

CHAPTER III

PROCEDURES

This chapter contains the restatement of the problem, the research design and purpose of the study. The chapter then continues with the derivation of hypotheses, the statement of the general hypothesis and their specific hypotheses. This is followed by subjects, sampling procedures and a description of the instruments used in the study. The variables are listed along with the descriptions and operational definitions of independent and dependent variables. Finally the chapter contains an explanation of the statistical approach to analyses of the data, including limitations of the study and a final summary.

Restatement of the Problem

Past experience in school reform has demonstrated that even when innovations are fully implemented they have little likelihood of being sustained (Fullan, 2001). Individuals involved in educational reform efforts recognize that there is a point where an innovation fully "routinizes" into the organization and becomes more than a model to be implemented; it is owned, adapted and sustained by the stakeholders within that organization (Rogers, 2003). There is scant literature that focuses on prerequisite behaviors and programmatic actions for sustainable school change, let alone the policy implications for school reform. Therefore, there is a need to understand the

programmatic path to sustained change so that future policy initiatives can build in the

necessary supports for long-term, positive change. The purpose of this investigation was

to test a theory-based model of sustained school change and the contributory constructs

for that model against the empirical evidence of Reading First Ohio.

Research Design

Ex post facto studies occur after the data collection and as the name states, refers

to any non-experimental research design in which the subjects are selected because they

have previously been exposed to a particular program, condition, or treatment (Kirk,

1995). This research study used an ex post facto research design with hypotheses and

tests of alternative hypotheses (Newman, Newman, Brown, & McNeely, 2006; Pedhazur

& Schmelkin, 1991). Although this design has inherent weaknesses, the validity of this

design was increased by stating relevant hypothetical research hypotheses. According to

Newman and Newman (1994), "ex post facto research with hypotheses and tests for

alternative hypotheses is considerably more powerful in terms of internal validity than

pre experimental, ex post facto designs with no hypotheses, and ex post facto designs

with hypotheses."(p. 112) Newman and Newman further stated that this type of research

design has a potential of higher external validity when compared to quasi and true

experimental designs when conducted under non-random sampling conditions.

Kerlinger and Blee (2000) identified three weaknesses of ex post facto research

design. The weaknesses include: (a) the inability to manipulate the independent variable,

(b) the lack of power to randomize, and (c) the risk of improper interpretation. When the

independent variables cannot be manipulated due to ethical or convenience reasons it

limits interpretation to the demonstration of relationships and excludes interpretation of

causation (Kazdin, 1995). However, since the current study focused on testing a model against the empirical evidence of Reading First Ohio and is not intended as a study to explain causal effects or to generalize causal relationships, the need to manipulate any independent variable is minimized as is the impact of randomization.

Derivation of the General Model and Research Hypothesis

The general research hypothesis is drawn from the body of literature presented in the Review of Literature (Chapter II) and in the Reading First Program Model itself (USDE, 2002). The Reading First Ohio model is built upon research concerning best practices in early literacy and the components of implementation, while elements of the model are drawn from those variables known to be implemented and the constructs that individual implementation variables form as a collective.

The variables identified reflect the body of literature for school change and the constructs or latent variables that are reflective of the major vehicles for change. A Structural Equation Model (SEM) consists of latent variables or factors that are not measured directly but are operationally defined by the researcher and linked to observable measurable variables. Latent variables are identified as either endogenous or exogenous. Exogenous variables are influenced by variables outside of the model itself, while endogenous variables are influenced by the exogenous variables through the specified model. Fluctuations in the endogenous variables are assumed to be explained by the model and the theory underlying the model. (Osborne, 2008)

Structural Equation Model representation.

SEM uses specific conventions for the representation of the model. Exogenous variables are represented with squares. Endogenous variables are represented with

circles. Direct influence of one variable on another is notated with a path. Paths are represented with arrows: the arrow direction points from the effected variable to the input variable. Error as a value is associated with each of the variables and factors, reflecting random measurement error, and is represented in the theoretic model. SEM allows the researcher to fix the value at one in the model specification phase. (Osborne, 2008) The model development phase of SEM involves the researcher identifying the variables of interest based upon the literature and diagramming the theoretic model to be tested against the data.

Exogenous variables in the proposed model.

Four exogenous latent variables are proposed; Leadership (LEAD), Classroom impacts (CLASS), Model Fidelity (MODEL) and Challenges (CH).

Leadership. Multiple researchers have acknowledged the importance of involved leadership in any sustainable change process (Bryk et al, 1998, Fullan, 2001, Elmore 2000) In Reading First Ohio, principals were trained in Classroom Walkthroughs (Elmore, 2000; Salzman 2006) and in the use of data-driven decision making (Fullan, 2001). The exogenous latent variable labeled leadership (LEAD) will be measured with the following variables; Walkthrough frequency per month (CWTfreq/mth), Attendance at grade level and/or team meetings per month (ATTEND).

Classroom impact. Due to the restrictions inherent in collecting data at the teacher or classroom level (as described in the sample section), the latent variable classroom impact encompasses teacher, student and the context of the learning environment. Classroom improvements in the context and support of teaching as measured by the Early Language and Literacy Classroom Observation (ELLCO) and

Survey of Enacted Curriculum (SEC) are validated by the work of Smith & Dickenson. (2002) Studies have shown that student achievement can be a reflection of improved classroom instruction (Fullan, 1993; Guskey & Huberman, 1995; Newman & Wehledge, 1995). The exogenous latent variable labeled classroom impact (CLASS) is measured by student achievement in the form of DIBELS (DIB), Terra Nova (TN) and the Ohio Achievement Test (OAT). Classroom context is measured by the ELLCO as a reflection of a literacy supportive classroom.

Model fidelity.

The full implementation of the initiative is an absolute prerequisite to intent or ability to sustain that program beyond the life of the grant. (Berman & McLaughlin, 1977; Fullan, 2001; Datnow & Stringfield, 2000) RFO built specific accountability for program implementation into the initiative to insure that each site was implementing with fidelity within a relatively short time. This variable (MODFI) will be measured in two ways; the time taken to reach a basic but full implementation in months (IMP), and the number of months that individuals remained fully implementing (TOTFULLPM). Additionally the measure instituted by the grant itself definition of progress (DP) will be incorporated as a measure demonstrating the overall boundaries and compliances required by the Reading First grant itself.

Challenges to sustainability.

During the course of implementation of the initiative there were significant hurdles to full implementation that acted as a limiting factor on ability to implement and ultimately sustain. These included building construction, building relocation and/or redistricting (CON), and/or administrative changes (AD) which served to shift the policy

focus and support internally. One additional variable was included here based on the experience of the implementation; Layers (LYR). The layers variable reflects the number of decision making bureaucratic layers between the District Coordinator (as the leader of the project) and the Superintendent of the District (as the major policy maker for the district). It is expected that this factor, Challenges (CH) and the associated path will operate in a negative direction. Since these variables could clearly be identified and measured it was determined to be prudent not to allow this negative variability to shift to the error terms. (Datnow & Stringfield, 2000)

Endogenous variables in the proposed model.

Two endogenous latent variables are proposed; Adaptive Collaboration (AC) and Ownership (OWN). These two variables are derived from the research that states that collaboration and ownership are prerequisites to moving into sustainability.

Adaptive collaboration.

Adaptive Collaboration is a process where individuals within the organization meet on a regular basis with the specific intent of reviewing the program and student data and to make collaborative decisions around those issues. Adaptive Collaboration as an exogenous latent construct is created by the individual inputs of each of the three positive exogenous constructs and by the one proposed negatively contributing construct. Hargreaves and Fullan (1998) emphasized the importance of assessment literacy for teachers and the importance of working collaboratively to review student data to make instructional decisions. Newmann and Wehlage (1995) also talked about the importance of professional learning communities in supporting both teaching and instruction.

52

Ownership.

Ownership is the highest level of commitment and in this study serves as both an indicator of and proxy for sustainability. This is the level that represents Rogers (2003) level of "routinization" within the organization. Ownership is indicated by actions within the organization that require multiple levels of commitment. As a construct Ownership builds on the previous constructs but it also extends beyond the scope of the previous. Ownership is directly measured by the degree to which a building has elected to extend their work beyond the original program parameters (EXTRFO). Ownership is indicated by a spreading of the program processes and/or content to additional grade levels, to new buildings or spreading for processes to new content areas. Ownership is the product of group commitment and a focused change effort. (Coburn, 2003; Fullan, 2001).

Participants

Participants in this study were from participating districts in the Reading First Ohio initiative during the implementation years of 2003 to 2009. Based on the grant criteria established by the U.S. Department of Education, and operationalized by each state, all participating districts were low performing and high poverty. Identification of the buildings that would participate was a district decision, based on selection criteria, but once selected all K-3 classrooms within an identified building were required to participate.

Sampling Procedures/Strategies

Schools were invited to participate in Reading First Ohio based upon their poverty levels and student performance. United States Department of Education and Ohio Department of Education guidelines determined eligibility for participation. Poverty

levels were designated as 15% or more children from families with incomes below the poverty line, 15% or greater "free and reduced lunch," or 6,500 children from such families. Districts not passing 60% or more of the students on the Ohio Department of Education Grade 4 Reading Proficiency test in 2003 were identified as having the highest percentages of students reading below grade level or as low performing (ODE, 2002). Eligible districts were rank ordered from highest poverty and lowest performance academically. Initially, based on these criteria a target applicant pool of 26 districts was created and of these districts, 20 eventually applied for and participated in Reading First. An additional group of districts was identified and 11 districts from the second identification applied. A total of 31 districts participated in Reading First Ohio. District participation was voluntary and building participation was determined by the same selection criteria as the district participation, and required sign-off by the teacher's union and, where appropriate, the administrative union. All kindergarten through grade three teachers and students in the selected schools were included in this study. In addition to the teachers and students, the building administrator, and the required Reading First staffing (Literacy Specialist per building, Data Manager, Resource Coordinator and District Coordinator) were all included in data collection.

Instruments

The instruments chosen for this study were selected by Reading First Ohio or created by Reading First Ohio Center in cooperation with the Ohio Department of Education, Office of Reading First and the U.S. Department of Education, Reading First Center. Fidelity of program implementation was measured by the Program Monitoring Tool (PMT). The student achievement data were collected with three different

instruments: the Dynamic Indicators for Basic Early Literacy Skills (DIBELS), the Terra Nova (TN), and the Ohio Achievement Test (OAT). Data about teachers, data managers, literacy specialist, and principals were collected through end-of-year surveys conducted as part of the information gathering for Reading First database updates and program review. Changes in classroom practices were collected using the Early Language and Literacy Classroom Observation Instrument (ELLCO) and by the Survey of Enacted Curriculum (SEC).

Program Monitoring Tool.

The Reading First Program Monitoring Tool (RFPMT) is a rubric style instrument designed to be used in conjunction with a document review process. The instrument was developed specifically for Reading First Ohio (Newman & Salzman, 2003) as a means of evaluating program implementation in accordance with Ohio's application to the United States Department of Education. There are two dimensions specified for implementation evaluation in the grant application; compliance with the three standards (and 18 indicators), and to benchmark the progress that districts and schools make toward fully implementing Reading First Ohio with fidelity. The three standards are based upon the knowledge bases of scientifically-based reading research (SBRR) (National Reading Panel, 2000; National Research Council, 1998) and professional development (National Staff Development Council, 2001). The program standards are: Professional Learning, Comprehensive and Coherent Program Design, and Systematic and Explicit Reading Instruction. Estimates of content validity were established through review by the State Reading First staff as well as the staff at the Central Regional Reading First Technical Assistance Center (CRRFTAC) at the University of Texas in November of 2003. Minor

revisions were made prior to the piloting of the instrument. Finally following the pilot, the instrument and protocols were further reviewed by a panel consisting of the State program leadership and field personnel to finalize the instrument.   Reviewers were trained in the use of the instrument and practiced until inter-rater reliability exceeded .80 (Reading First Ohio, 2003).

Program monitoring was conducted by the trained personnel who reviewed the program documentation provided by building personnel and rated the documentation as a proxy for the implementation of the underlying program on a three point scale. A total of 54 points was possible, with a score of 49 out of 54 necessary for continuation funding at the end of year 2. Buildings were required to submit to Program Monitoring three times per year. Scoring was recorded and written feedback as well as technical assistance was provided based on the findings. Program monitoring process and reporting formed the foundation of technical assistance.

> The Program Monitoring process was designed to fulfill multiple purposes
> for the Ohio Department of Education, the Reading First-Ohio Center, and
> the recipient school districts themselves. These purposes are: Self
> Assessment, Compliance, Documentation, Measure of Implementation and
> Technical Assistance Jump Start (Salzman, 2008).

Through the systematic application of technical assistance all schools in the program were able to reach minimum scoring on Program Monitoring to allow continuation of funding.

Dynamic Indicators for Basic English Literacy.

Dynamic Indicators for Basic English Literacy (DIBELS) was created by Good and Kaminski (2002). This test is used to assess measure the acquisition of early literacy skills from kindergarten through sixth grade. This test was given to each student individually three times per year within a 2-week testing window. There are four basic developmental skills that this instrument assesses: Initial Sound Fluency (ISF), Phonemic Sound Fluency (PSF), Nonsense Word Fluency (NWF), and Oral Reading Fluency (ORF). The single probe reliability for the ISF ranged from a low of .61 to a high of .86 and was only used in Kindergarten. PSF had a reliability of .74 in kindergarten. In kindergarten the NOW ranged from a low of .86 to a high of .94. In first grade the NWF as about the same with a reliability ranging from .83 to .94. Oral Reading Fluency (ORF) started in first grade and continued through third grade. The lowest reported reliability was .92 with the highest equal to .97. On ORF test-retest reliability was .97. In addition both predictive and concurrent validity was conducted on ORF. For predictive validity values ranged from a low of .62 to a high of .72. Concurrent validity ranged from a low of .67 to a high of .82, thus suggesting that this instrument is both valid and reliable (Dynamic Measurement Group, 2008).

Terra Nova.

Terra Nova (TN) was developed to provide achievement scores that are valid for most types of educational decision making. The test results include measurement of achievement for individual students related to a current national normative group. Progress can be tracked over years and across grades. Terra Nova can also be used in a criterion-referenced manner to measure gains in student academic strengths as well as to

identify weaknesses in each of the content areas. This test can be used administratively to make programmatic decisions and assess overall class progress. Content validity was established by expert judges who compared Terra Nova content with current classroom practices and curriculum nationally. These expert judges stated that the assessment accurately represents the important educational objectives seen throughout the nation. The construct validity was approximated by reviewing the correlations between Terra Nova CTBS complete battery and TCS/2. The Reading Composite subscale and the other tests correlations ranged from .56 to .80 with a total TCS/2 correlation of .72.

Ohio Achievement Test.

The Ohio Achievement Test (OAT) is a criterion referenced test that was created by the Ohio Department of Education to assess mastery of state academic content standards. This test is first administered to students at the third grade level and therefore only data from third graders in RFO schools was collected. The validity of the OAT test was expert judge done through committee. The validity of the OAT was expert judgment done through committee. (Personal communication with Paula Mahaley and Chad Richardson Data Manager, Office of Literacy Center for Curriculum and Assessment Ohio Department of Education). There is a yearly report on the reliability of the OAT produced by the Ohio Department of Education. From the onset of this instrument the reliability has ranged from .86 to .92. The 2008 reliability was reported at a .90 (Office of Assessment, Ohio Department of Education).

Survey of Enacted Curriculum.

Survey of Enacted Curriculum (SEC) was created by Wisconsin Center for Educational Research (WCER) in 1995. The SEC is a reliable data collection tool that

58

provides an objective method for analyzing the degree of alignment between instruction and state content standards. This survey is a self-reported on-line survey. Teachers at the end of the school year have a 3-week window to log on and reflect on their teaching practices for that year (Blank, R 2002). The reliability and validity for both the English Language Arts and the Math/Science section of the SEC were not well reported. There were several more studies that investigated the Math/Science section of the SEC since it was this instrument's original focus. The English Language Arts section was not development until 2002 and the standards were not mapped until 2003. There were expert judges that worked with the WCER and the Ohio Department of Education on aligning Ohio state standards to the SEC questions. There was no reported internal reliability, test-retest or predictive validity estimates available for this instrument. All reports indicate that there is high reliability and validity but do not report any numbers. This conclusion was reached following contact with Chris Woolard, director of the SEC project for ODE, Learning Points Associates, and John Smithson, Director of the SEC online and the WCER.

Early Language and Literacy Classroom Observation.

Early Language and Literacy Classroom Observation (ELLCO) was created by Educational Development Center (2002). This observational field-test was designed to assess the effectiveness of professional development and teacher practices. Trained observers completed the three components of Literacy Environment Checklist, Classroom Observation with Teacher Interviews and Literacy Activities Rating Scale.   This study utilized the Classroom Observations scoring as an indicator of classroom implementation and best practices. Identified teachers from kindergarten through third grade were

observed in the fall and again in the spring. Scores were aggregated by grade level or by

building per practices agreed upon as part of the grant administration. The items that

created the subscale of Classroom Observation resulted in a Cronbach's Alpha of .90

which indicated very strong internal consistency. This subscale also showed moderate to

high correlations to all of the other subscales ($r = .034$ to $r = 0.65$) (Smith & Dickinson,

2000).

Structural Equation Modeling software.

In SEM, the software used to conduct the analysis has an impact on the analysis

itself. Mueller and Hancock (2008) stated that the best practices in SEM writing should

"include a reference to the SEM software package used since results can vary not only

across programs, but also across versions of a single package (mainly due to differences

and continual refinements in estimation algorithms)" (p. 504). There are multiple

software applications of SEM offering slight differences in modeling and parameter

estimation defaults; LISREL (Joerskog & Sorbom, 2006), EQS (Bentler, 2006) and

Mplus (Muthen & Muthen, 2006).   Analysis Moment Structures version 17 (AMOS 17)

published by SPSS developed by Arbuckle (2007) was used for this study. It is a

Windows-based SEM software that has three modes of model specification; AMOS

graphics, AMOS VB.NET and AMOS C#. The choice of mode is a researcher

preference as there is no substantive difference between the modes since the underlying

algorithms do not vary by mode (Mueller & Hancock, 2008).

Variable List

The variables tested within the model or as part of this study are included in Table 1

below:

Table 1.

Variables, Description and Measurement Level

Variable Name	Variable Description	Measurement
Bureaucratic Layers (LYR)	In small districts grant personnel have more direct access to policy decision makers. This access facilitates grant administration and communication about the grant. In some districts the grant administrator reports directly to the superintendent and in others there are as many as 6 layers of personnel in the organizational chart between the superintendent and the grant administrator.	Interval Ratio
Construction/Redistrict (CON)	The Ohio Department of Education had a major school renovation and reconstruction project during the years of Reading First Ohio. Many of the districts experienced one or more building projects and as a result had to displace school faculty and students. This disruption is a possible barrier to a quality implementation.	Interval Ratio
DIBELS (DIB)	Dynamic Indicators of Basic Literacy measures student literacy data collected in grades K-2 as part of the ongoing programmatic process.	Interval Ratio

Variable Name	Variable Description	Measurement
% Superintendent Change (%SuptChng) % Principal Change (%PChng)	This variable represents the possible changes in leadership that are common in public school districts and can influence school reform efforts.	Interval Ratio
Program Extension (EXTENDRFO)	This study proposed that schools who are most invested in the school reform effort will extend the program beyond the programmatic boundaries. In the case of Reading First Ohio, some schools elected to extend their reform efforts to grades 4 and beyond.	Interval Ratio
Months to reach full implementation (IMPPM)	Reading First Ohio districts are monitored three times each year to determine the fidelity of implementation. This variable is an indication of the number of months the building took to reach a full, yet basic, level of implementation.	Interval Ratio
OAT (OATMean)	Grade 3 Ohio Achievement test in Reading scores by building.	Interval Ratio
Percentage of Teachers Fully Implementing (%Teacher Implement)	Percentage of K-3 teachers reported as fully implementing the basic components of a reading program as outlined in the Reading First grant.	Interval Ratio

Variable Name	Variable Description	Measurement
%Principal Support (%PRINSUPPT)	District Coordinator report of principal participation, support and demonstrated activity in the program.	Interval Ratio
PMT Months of Full Implementation (TOTFULLPM)	This variable represents the number of consecutive months of full implementation.	Interval Ratio
Survey of Enacted Curriculum (SEC)	This variable represents classroom implementation of instructional and standards-based practices learned through professional development.	Interval Ratio
Terra Nova (TNGrowthMean)	This variable represents the Terra Nova scores of the third graders in Reading First Ohio aggregated by building as a mean growth score.	Interval Ratio
Walkthrough Frequency per month (FreqCWT/qtr)	This variable represents the number of principal walkthroughs in the building and is an indicator of principal involvement and of commitment to program fidelity measured quarterly.	Interval Ratio

Data Collection Procedures

The data used in this study are the Reading First Ohio (RFO) data that has been collected across the 6 years, by the Reading First Ohio Center or through the Ohio Department of Education. Every student, teacher, literacy specialist, data manager,

resource coordinator and principal that worked in a Reading First Ohio school during this time is included in this study through different data paths.

Schools were invited to participate in Reading First Ohio based upon their poverty levels and student performance. A total of 31 districts participated, 145 buildings inclusive. Each district signed the Ohio Department of Education Reading First Ohio Grant Assurances (ODE, 2003). These assurances specified guidelines for ongoing participation. Every district agreed comply with the implementation requirements specified by the federal grant and by Reading First Ohio guidelines. Every district agreed to collect ongoing data about the embedded professional development. In addition, four times a year the schools are required to send the Ohio Department of Education student test scores. One hundred percent compliance is expected or the district risks loss of funding.

Every effort was made to preserve the confidentiality of participants in the evaluation process. To protect confidentiality of the students the state student identifier (SSID) was used in place of names. This SSID is a number that follows the student anywhere within the state. This allowed students to be tracked if they moved from one RFO district to another RFO district. For teachers the issue was more complex. Ohio as a local control and unionized state had to work within union agreements for teacher confidentiality. No teacher level/classroom level data could be reported in a non-aggregated form. Due to this, all teacher level data were aggregated by grade for each of the 145 RFO schools to insure teacher confidentiality. Grade level by building data is the closest proximity to classroom that is reported.

Also as a condition of participation in Reading First Ohio, each building was required to prepare program monitoring evidence collection three times during the year. The evidence collection and review was guided by and scored according to the Program Monitoring Tool (Salzman & Newman, 2003)   The collection of supporting evidence was examined by trained reviewers and a feedback score was provided to the building administrator and Reading First staff at the building and district level. Feedback from the program monitoring process was used to provide technical assistance in program implementation as well as to monitor the progress of the schools toward full implementation of the initiative. In year two of the building participation a minimum score on program monitoring and a cut score on student achievement at Grade 3 determined continued funding; this was called the definition of progress. Attainment of a minimum score on program monitoring along with a ten percent reduction in students who were below 50% on the Terra Nova at the end of Grade 3. Terra Nova was the selected measure as a bridge measure during the development and validation of the Ohio Achievement Test for Grade 3.

> District programs that show no progress the first year will be allowed to continue for one more year with intensive assistance. At the end of the two years, the district program that shows no progress will be discontinued. Districts that show minimum progress in the first year will be allowed to continue with targeted technical assistance intervention. Districts showing strong evidence of efficacy and student achievement will be eligible for enhancement funding (USDE RF, 2002, p. 70)

Definition of progress was defined as an indicator of satisfactory progress toward full and active implementation and therefore the failure to reach definition of progress was considered a failure to reach a level of implementation sufficient to impact student achievement levels.

Statistical Analysis

Descriptive analysis will be conducted to provide an overview of the data patterns and building characteristics both prior to initiating the Reading First Ohio model and at the end of fiscal year 2009. To understand the underlying structures within the school change sustainability model as demonstrated through Reading First Ohio, a theoretic model will be developed and identified using Structural Equation Modeling (SEM). Once the parameters have been identified and the model tested for fit, modifications will be made as necessary. Upon completion of the model development, the model will be further tested against the expert scoring as means of gaining further confirmation that the model functions in accordance with the intent of the Reading First Ohio model and school change efforts.

Structural Equation Model.

Structural Equation Modeling is the analysis method proposed because of its usefulness in examining a theory structure as a whole. Senge (1990) advocates systems thinking as the primary means for promoting a learning organization and for discouraging repeating the same mistakes that limit the long term growth of an organization. Too often studies have focused on individual input variables or have focused on single programmatic outcomes (student achievement, implementation etc) as the dependent variable in the study of school change. Many studies of school change have focused

exclusively on studying the effectiveness of a given school reform model or process. While those contribute to the body of literature, this study shifts away from specifics of the program and focuses on a process of developmental capacity building within elementary buildings in Ohio as they move toward sustained school change efforts. While Reading First Ohio provides the programmatic data for analysis, this study is not focused on the impact or outcomes of the programmatic model itself.

Mueller and Hancock (2008) identified SEM as not so much a statistical technique but as a modeling process that has several stages: initial model conceptualization, parameter identification and estimation, data-model fit assessment and potential model modification. Researchers are cautioned against using SEM as an exploratory tool but it is attractive for model and theory development because as a confirmatory method it provides a comprehensive manner for assessing theoretic model (Anderson & Gerbing, 1988). The model building aspects of SEM will allow this study to build from theory to practice and to contribute to the body of literature an empirical study of a school change model. One of the strengths of SEM is that it provides a strong connection between the underlying theories and the statistical analysis. SEM builds first from the theoretic model and then tests that model through analysis. In this way, SEM is a theory based analysis that confirms the proposed theory with the data collected. As an analysis tool, SEM reduces measurement error and provides a strong parameter estimate through the linking of path analysis and factor analysis. (Kaplan, 2009) Structural equations modeling (SEM) requires that the theoretic model be first specified based on theory and on the research knowledge base. Mueller and Hancock (2008) stated, "In the strict sense, any hypothesized model is, at best, only an approximation to reality; the

remaining question is one of degree of that misspecification" (p. 490). The initial step in

model specification is a thorough knowledge of the literature that forms the basis for the

model.
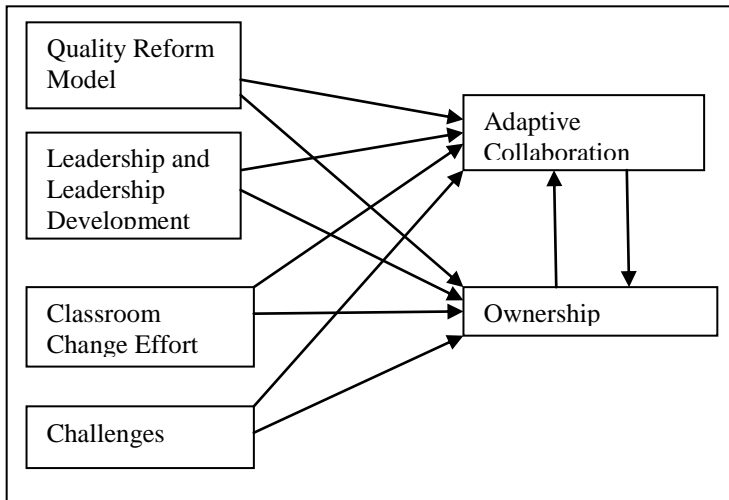
Figure 2.Proposed Theoretic Model

Methodological limitations of SEM.

The three stages of the SEM process as identified by Mueller and Hancock (2008)

include the final stage of model modification. Although ability to test a theory against

the data generated by practice is a strength of SEM, the model modification process can

be a weakness as well. At this point the researcher is able to adjust the model in an

attempt to improve the "fit" between data and theory. Although SEM provides an

opportunity to review the underlying data structures and supplement theory, there is also

the risk of moving away from the original theory base as the structural equation model is

modified to fit the data more closely. Spanos (1986) cautioned researchers against the

temptation of moving the model too far away from the theoretic and literature based

structure to that of the data generated by the study. The process of model modification

must be guided first by the theory base and secondly by the SEM model fit. Moving too

far from the original theory base threatens to create an illogical model, regardless of the statistical fit. Poor fit even after reasonable modification is indicative of a poorly constructed theoretic model or inadequacies in the theory base and requires a renewed review of the existing literature or the generation of new research. SEM is one of the few statistical applications where the results are likely to show that the original theoretic model is in error. "If it is true that a proposed model does not reflect reality, then reaching a conclusion of misfit between data and model should be a desirable goal, not one to be avoided by careless respecifications until satisfactory levels of fit are achieved" (Mueller & Hancock, 2008, p 506). Joreskog (1993) specified three possible SEM strategies; strictly confirmatory (SC), alternative models (AM) and model generating (MG). The MG is the most common method demonstrated in the body of SEM literature (Byrne, 2010). This study is focused on the application of SEM as a tool in confirming the alignment of theory and the reality of the experience of RFO schools. It is proposed that the approach be model generating (MG) to allow for minor model respecifications in the path structures.

SEM and sample size.

Traditionally SEM is considered a large sample analysis with requirements of normal distribution and preferably randomization. Originally researchers were encouraged to use "rule of thumb" estimates on sample size. Bentler and Chou (1987) stated that the minimum sample size for SEM ranges from 5 to 10 cases per estimated model parameter depending on the estimation method used. More recently, researchers have focused on determining the degree to which SEM is robust under circumstances of non-normality or small sample sizes. Multiple Monte Carlo studies have been conducted

to review the degree to which parameter estimates can be reliable and found that corrections can be made for small sample size and nonnormality (Fan et al 1999; Herzog et al 2007; Jackson, 2003) Nevitt and Hancock (2004) findings suggested that sample size adequacy is best measured by the ratio of subjects-to-estimated parameters (n:q) rather than an absolute sample size. Multiple researchers have examined means of estimating model parameters (Browne 1982, 1984, Satorra & Bentler, 2001; Yuan & Bentler, 1997). Nevitt and Hancock (2004) found that the Bartlett k-factor correction to maximum likelihood estimates yielded and empirical power at 81-100%. The Satorra and Bentler statistic (T_{SB}) yielded model rejection rates of 90% at n = 35. Ultimately following a study of multiple methods of parameter estimation with corrections for non-normal and small sample sizes, the authors determined that researchers can effectively model data using SEM in cases of small to moderate sample sizes. They also found that the statistic was robust even with severely nonnormal data and samples sizes of less than 100. Some of the fit statistics even performed well with non-normal data and a sample size of only 50. Ultimately, the authors recommend that researchers obtain samples sizes large enough that n:q is less than or equal to 2:1. The authors also recommended the T_{SB} statistic for assessing overall data-model fit with small samples. AMOS 17 software contains multiple corrections for sample size concerns.

Expert judgment.

One of the limitations of SEM is the possibility of creating a statistical model that while mathematically possible given the data and theoretically plausible given the research, is not in alignment with reality. Upon completion of the analysis using SEM, the model will be reviewed and beta weights for each path examined by district.

70

Individual district paths will be compared to outcomes identified by expert judges working with the districts. As part of the RFO technical assistance monitoring process, Regional Consultants (expert judges) were asked to place districts along a continuum ranging from non-compliant through basic grant implementation to demonstrating authentic ownership and ready to sustain.   Relative performance of each district can be reviewed as part of an overall confirmation of the validity of the SEM. Through this process the SEM will be reviewed for its alignment to the reality of the program implementation and contextual issues of the district.

Seven regional consultants engaged in the examination of district performance as part of the RFO technical assistance planning process. The seven Regional Consultants were fully trained in the program monitoring process and had each engaged in excess of 100 work days of technical assistance to the districts and buildings. Reliability of their program review was established as part of the program monitoring process (see Instrumentation).

The path estimates derived from the SEM was used in a regression equation to create an ownership score for each participant district. Districts were rank ordered based on the derived ownership score and the scores correlated using a Kendall's Tau to the scores obtained through the expert judges.

Database Preparation

Data were collected from databases at the Ohio Department of Education and the Reading First Ohio Center. The archival databases were entered into SPSS version 18 (PASW 18) and merged. The PASW data file was then merged into the AMOS analysis. The unit of analysis for this study is all elementary buildings with unique IRN

identification participating in Reading First Ohio; N=145 records. The sample had

missing values distributed across multiple variables. Variables were removed from the

study as necessary given missing data. No method of data imputation was considered

given the large percentage of missing values in specific variables (Little and Rubin,

1987).

Reading First Ohio Center individual student outcome data from the two

assessments, DIBELS and Terra Nova, were aggregated by school IRN as mean growth

across years. The variable title "Full implementation" was calculated based upon each

individual Program Monitoring for each individual school IRN based upon the grant-

specified, annual schedule of program monitoring for each site. Full implementation for

the purpose of program monitoring during the program initiation was determined to be a

score of 53 or 54 out of a possible 54 on the program monitoring tool. TotalFullPM as a

variable is an aggregate score based on the number of program monitoring periods where

the site was able to maintain a score of either 53 or 54. All other records were extracted

from the database and identified by school IRN.

One of the underlying assumptions of multivariate analysis and structural

equation modeling is distribution normality (Byrne, 2001; Hair et al. 1998, Kline, 2005)

Data were screened for distribution. There were no outliers and the residuals in the

analyses were normally distributed so no transformations were required.

Characteristics of the Sample

Reading First Ohio Participants in this study were from participating districts in

the Reading First Ohio initiative during the implementation years of 2003 to 2009.

Based on the grant criteria established by the US Department of Education, all

participating districts were low performing and high poverty. The State of Ohio identified

eligible districts for participation. The identification of the buildings that would

participate was a district decision, based on selection criteria, but once selected all K-3

classrooms within an identified building had to participate. Poverty levels were

designated as 15% or more children from families with incomes below the poverty line,

15% or greater "free and reduced lunch" or 6500 children from such families.   Districts

not passing 60% or more of the students on the Ohio Department of Education Grade 4

Reading Proficiency test were identified as having the highest percentages of students

reading below grade level or as low performing.(USDE, 2002) A total of 31 districts

participated in Reading First Ohio, with a total of 145 separate building IRNs within

those districts. District participation was voluntary. All kindergarten through grade three

teachers and students in the selected schools were included in this study. In addition to

the teachers and students, the building administrator, and the required Reading First

staffing (Literacy Specialist per building, Data Manager, Resource Coordinator and

District Coordinator) were all included in data collection. (ODE, 2002).

Demographics of the Study Population

The participating districts in Reading First Ohio are by definition high poverty and

low achieving as measured by state testing. The student population in participating

elementary schools and grade levels is 71.6% economically disadvantaged. The student

enrollment is 46.2% African American, 41.7% White, 6.5% Hispanic and an additional

5.1% is reported as mixed race (see Table 2).

Table 2

Demographic Statistics on the Student Enrollment

	Frequency	Percent	Valid Percent	Cumulative Percent
Gender				
Female	30865	48.7	48.7	48.7
Male	32559	51.3	51.3	100.0
Ethnicity				
Asian	246	0.4	0.4	0.4
African	29308	46.2	46.2	46.6
American				
Hispanic	4109	6.5	6.5	53.1
Indian	104	0.2	0.2	53.2
Mixed	3230	5.1	5.1	58.3
White	26426	41.7	41.7	100.0
Economically				
Disadvantaged				
Not	17568	27.7	28.4	28.4
Is	44187	69.7	71.6	100.0

Urban/Rural Distribution

Ohio Department of Education classifies public school districts in Ohio by the type of community served. Reading First Ohio as a statewide initiative served 57.8% "Major

Urban-Very high poverty" districts, 24.5 % "Urban-Low median income" districts,

16.3% "Rural/agricultural – High poverty, low median income" districts (see Table 3).

Table 3

Urban/Rural Designation for Reading First Ohio Elementary Schools

	Frequency	Percent	Valid Percent	Cumulative Percent
Major Urban-Very High Poverty	85	57.8	58.6	58.6
Urban – Low Median Income, High poverty	36	24.5	24.7	83.4
Rural Agricultural – High Poverty, Low Median Income	24	16.3	16.6	100.0

Summary

This chapter provided an overview of the methodology used to implement this study. The procedures and data collection of the study were described along with the instrumentations used. The ex post facto research design was explained, and the population sample was described. A variable list and the derivation of general research hypotheses and statements of general and specific hypotheses were provided. In addition, operational definitions were described and limitations of the study were also listed.

CHAPTER IV

ANALYSIS AND PRESENTATION OF FINDINGS

Introduction

This chapter presents the results of this study. The first section contains the sample descriptive statistics and correlations among variables. The second section provides the factor analysis that forms the basis for creating the measurement model as well as justification for initial variable trimming. The third section establishes and tests the full structural equation model including the two components; the measurement models and the structural model. The final section will review the research questions posed in this study.

Descriptive Analysis of Variables

The descriptive statistics for the study variables are presented in Table 3. The study variables are those that are contained within the final structural equation model. It is important to note that in Table 3 Minimum values in Terra Nova Mean Growth (TNGM) are negative reflecting a below benchmark performance value for a given elementary school (see Table 4).

Table 4

Descriptive Statistics for Study Variables

	N	Mean	SD	Minimum.	Maximum
ELLCO	142	53.50	8.86	23.11	68.04
EXTRFO	141	1.23	1.61	00.0	5.00
FREQCWT	145	1.738	3.73	00.0	20.00
OAT	131	400.42	10.40	372.00	424.25
%BLDGIMP	145	85.79	16.61	50.00	100.00
%PSUPPT	145	78.24	22.70	00.0	100.00
%TIMP	145	80.14	20.32	50.00	100.00
SEC	138	.1961	.0164	.14	.27
TNGM	125	1.22	5.57	-13.86	15.22
TFullPM	139	5.91	3.63	00.0	12.00

Model Development

This study focuses on developing the structural model for a single dataset and establishing initial baseline indicators for models of ownership and sustainability in school change efforts. The purpose of the study is to test the more global theories of sustainability of school reform efforts against the empirical experience of Reading First Ohio. Structural equation modeling is generally viewed as a confirmatory process seeking to apply goodness-of-fit measures to determine if the variance-covariance matrices are consistent with the theoretical path model proposed by the researcher and based within the literature. (Byrne, 2001) SEM consists of two phases; the measurement model and the structural model. The measurement model consists primarily of factor

analysis to demonstrate that the observed variables load to or build the constructs or latent variables. The structural model consists primarily of path analysis with latent variables. The process begins with a theoretic model that is based in the literature, proceeds to test the measurement model and then finally tests the structural model. Goodness-of-fit indices are then used to determine modifications that are necessary to improve the overall fit of the model with the data.

Refinement of the Proposed Theoretic Model

The theoretic model proposed (see Figure 1) and detailed previously was based on the prevailing theories of sustainability in school reform (Colburn, 2005, Fullan, 1999) and on the Reading First Ohio work plan (RFO, 2003). Upon review of the archival data and early analysis, three alterations to the model were clearly indicated.

It was determined that "Adaptive Collaboration" was not documented in a manner that would permit review of a unique impact from the process of adapting the grant collaboratively. This variable required information pertaining to the collaborative processes in place within the school. Little data were available to document the content of meetings although the meetings themselves were documented. Insufficient data were available to make it possible to incorporate this variable. This was executed at the time of data preparation and prior to initiation of analysis.
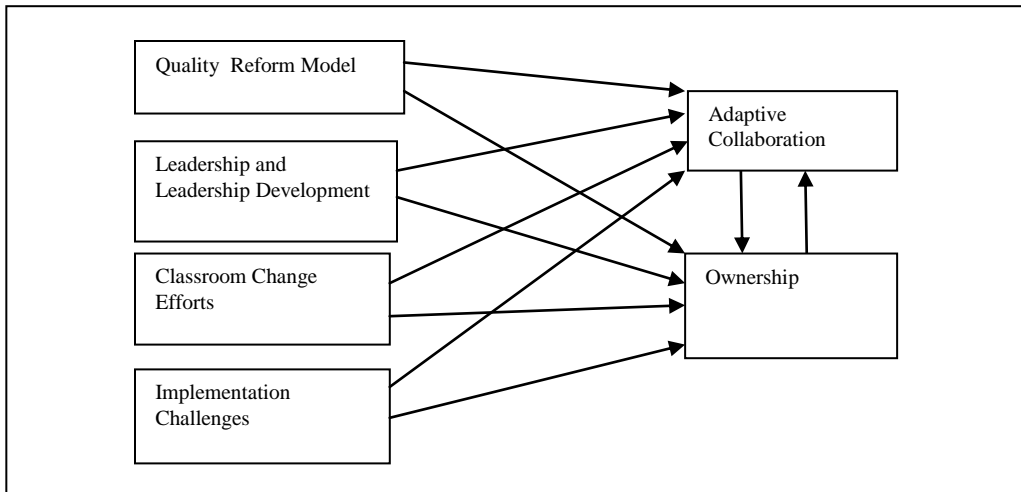
Figure 3.Proposed Theoretic Model

The theoretical model was revised in a manner that incorporated collaborative processes into the area of fidelity of implementation. In addition, review of the data demonstrated that the Reading First Ohio database was limited in the data collected concerning leadership. Early analysis demonstrated that it might be appropriate to allow the variables previously considered under the "Quality Reform Model" to load into either the "Classroom Impact" or "Leadership" areas. This was completed during the early factor analysis and measurement model phase.

Further, the area of implementation challenges did not clearly contribute to model development in the process of respecification of the model in the measurement model phase of model development. This construct was also eventually removed and reviewed as a separate correlation analysis. Although, districts and elementary schools experienced a frequent leadership change and construction within the schools, it was not a significant (negatively) contributor to the model and as such was also removed (see Figure 4).
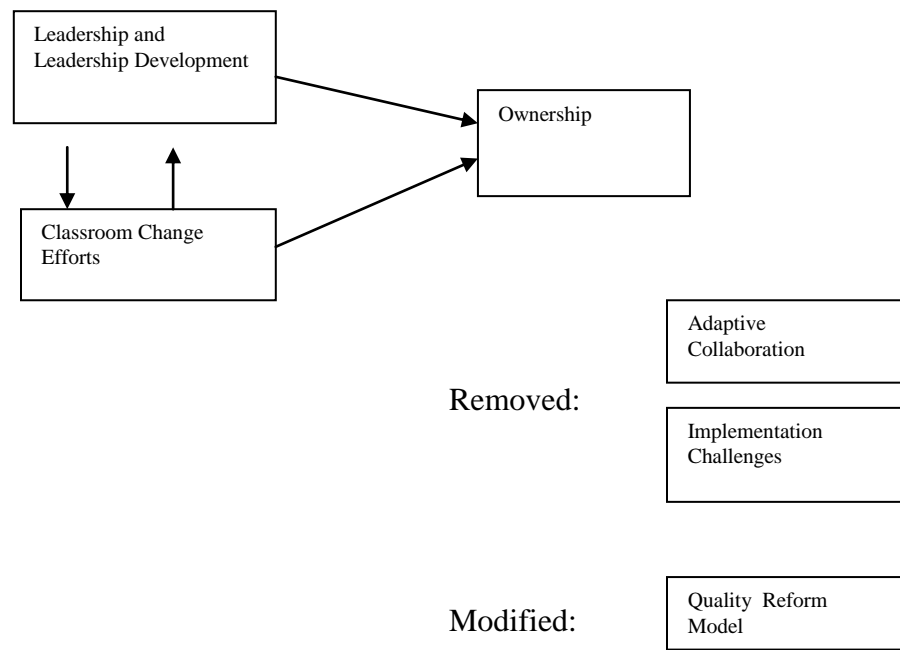
Figure 4.Revised Theoretic Model of Constructs

Although substantially more limited in scope, the model to be tested remains an

opportunity to begin to build a body of empirical literature around the construct of

ownership in change processes and specifically in sustained school change efforts.

Structural Equation Modeling Process

    Overview of the process.

Following the review of the data, an exploratory principal components analysis was

performed to review the underlying factor structure and to identify the viability of the

proposed theoretic model. The factor analysis was conducted using PASW. (see Table 4)

At this stage either factor loadings or the correlation matrix provides adequate

opportunity to review the data. Factor analysis was selected because the Reading First

Ohio data were determined to have substantial intercorrelation that might make structure

identification more difficult. At this stage, this is exploratory only as the model has been

identified through the literature. The exploratory analysis provides the researcher with a

starting point for the development of the measurement models. Particularly in the case of

archival data that were collected for a different purpose than the study, it is necessary to do exploratory analysis.

Utilizing the approach advocated by Byrne (2001), during the measurement model phase, the measurement model for each latent construct was evaluated independent of the structural model. (See Tables 5 and 6) As part of the measurement model process the researcher conducts additional factor analysis with identified groups of variables and then tests the measurement model using AMOS 17. Individual variables were evaluated for loadings, covariance and misspecification through the modification indices. If indicated individual variables were removed from factors to increase the overall fit. In the case of a poor fit a hierarchical approach was applied to evaluate alternatives consistent with both the literature and the original program descriptions.

Finally and after establishing the measurement model for each latent construct, the entire structural model is evaluated for fit. In this study the two measurement models were tested then the structural model was tested using AMOS 17. The model was reviewed by evaluating the Chi-square statistic, degrees of freedom, minimum discrepancy function (CMIN/DF), the comparative fit index (CFI), the normed fit index (NFI) and the root mean square error of approximation (RMSEA) (Byrne, 2001, Kline, 2005; Hair et al, 1998) Modifications were made based on misspecification indices. The process of model building and trimming continued until the model of best fit was identified.

Correlation of variables.

Per American Psychological Association (APA, 2002) guidelines the correlation matrix along with means and standard deviations of study variables is presented in Table 5.

Table 5

Correlations of SEM Variables

		TNGrowthMean	OATMean	TotalFullPM	%BldgImp	%PrinSupp	%TeachImp	FreqCWT	ELLCO	SEC
TNGrowthMean	r	1								
	Sig. (2-tailed)									
	N	125								
OATMean	r	-.091	1							
	Sig. (2-tailed)	.315								
	N	124	131							
TotalFullPM	r	.125	.300	1						
	Sig. (2-tailed)	.174	.001							
	N	120	126	139						
%BldgImp	r	-.022	.295	.524	1					
	Sig. (2-tailed)	.807	.001	.000						
	N	125	131	139	145					
%PrinSupp	r	-.046	.247	.280	.505	1				
	Sig. (2-tailed)	.608	.004	.001	.000					
	N	125	131	139	145	145				
%TeachImp	r	-.087	.346	.399	.813	.415	1			
	Sig. (2-tailed)	.336	.000	.000	.000	.000				
	N	125	131	139	145	145	145			
FreqCWT	r	.054	.194	.447	.357	.162	.267	1		
	Sig. (2-tailed)	.547	.027	.000	.000	.051	.001			
	N	125	131	139	145	145	145	145		
ELLCO	r	-.022	-.072	-.036	.002	-.030	-.109	-.004	1	
	Sig. (2-tailed)	.809	.418	.678	.978	.726	.196	.959		
	N	124	130	137	142	142	142	142	142	
SEC	r	-.260	.260	-.047	.066	-.019	.135	.076	.050	1

Principal Components Factor Analysis with rotation was conducted to explore the underlying construct structure of the variables identified (see Chapter 3) as possibly contributing to the study. (see Table 6) Given that the data were collected as part of an initiative and not uniquely for this study some misfit was anticipated. The purpose of this exploratory factor analysis is information, not identification of constructs. The constructs to be tested have been identified based on theory this test is not searching the data for underlying constructs.

Initial loading presented an array of four factors that might fit the original proposed model. Variables with low factor loadings or loading with variable groupings inconsistent with the literature-based model were hierarchically removed as part of the specification process. It is important to note that within SEM, factor loadings are used as a guide to measurement model construction but not as an absolute determinant. Ultimately it is the model fit along with the underlying research that determines the degree to which a variable or the entire model is appropriately identified. The literature and model specified the logic of two dominant constructs; Classroom Impact and Leadership. These were tested through the measurement models.

Table 6

Summary of Exploratory Factor Analysis Factor Loadings Used to Inform Structural

Equation Model Specification and Respecification

Items	Rotated Factor Loadings			
	1	2	3	4
% Building Implement	**0.881**	0.088	0.363	0.02
% Teacher Implement	**0.869**	-0.126	0.123	-0.084
% Principal Support	**0.753**	-0.144	-0.158	-0.018
Total Full PM	**0.494**	-0.195	0.454	**0.422**
Freq CWT/qtr	0.249	-0.256	0.727	0.053
OAT Mean	0.234	**-0.637**	0.238	-0.061
% Principal Change	0.107	**0.753**	0.003	0.293
SEC	-0.041	-0.232	0.263	**-0.683**
% Supt. Change	-0.065	0.397	**0.72**	-0.099
ELLCO	-0.112	**0.605**	0.138	-0.123
TN Growth Mean	-0.121	-0.048	0.181	**0.811**
Eigen Values	2.851	1.937	1.301	1.284
% of Variance	23.761	16.141	10.838	10.700

Note: Factor loadings over .4 appears in bold

Measurement Model

Based on exploratory factor loadings, underlying literature and the Reading First

program design, measurement models were identified and tested. The measurement

models were analyzed and corrections made to improve the fit of the measurement model.

Leadership variables were limited in the dataset. Fidelity of program implementation in the form of the variables Building Implementation and Total Full were incorporated into the review of this construct to determine if the construct would converge. A final test of the variables to include was conducted using a Principal Component Analysis with varimax rotation with the variables Building Implementation, Principal Support, Total Full PMT, Freq CWT. The variables loaded onto a single factor (without rotation) with factor loading values exceeding .4 (see Table 7).

Table 7

Principal Component Analysis for the Latent Construct of Leadership

Items	Leadership
% Building Implementation	**.867**
Total Full PM	**.619**
% Principal Support	**.555**
Freq CWT/qtr	**.459**
Eigen Values	2.172
% of Variance	54.298

Note: Factor loadings over .4 appears in bold

Leadership demonstrated statistical significance (Chi square=11.551, df=2, p=.003) Other fit indices were a CMIN of 11.551, CMIN/DF = 5.776, CFI of .914, and RMSEA of .182. Individual paths were all significant (p<.001). All error paths were set to 1.0 (See Figure 5).
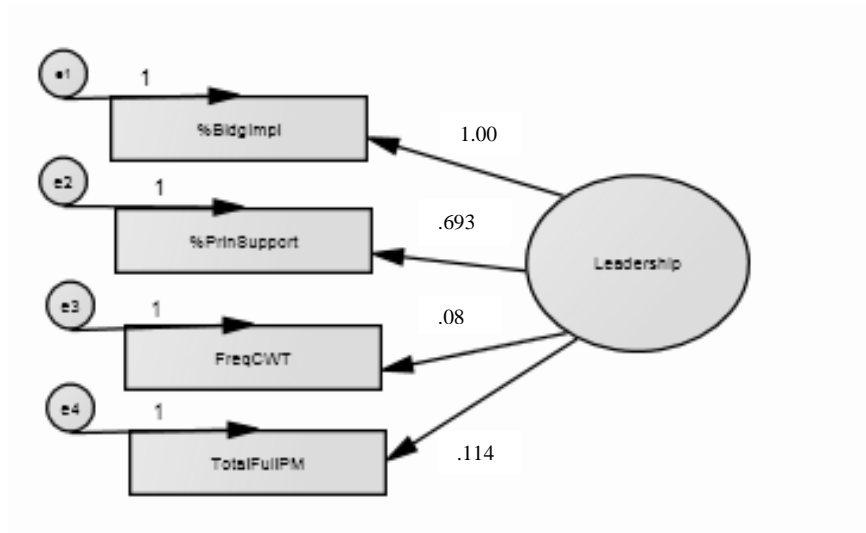
Figure 5.Leadership Measurement Model

Classroom Impact was a major focus of Reading First Ohio (2003) as well as strongly supported in the school change literature. As previously stated, the grant specified changes in classroom practices specifically designed to impact student achievement. A Principal Components Analysis with varimax rotation was conducted to identify the underlying structure of the classroom impact variable. The following variables had the potential of contributing to this variable; DIBELS Growth, TN Growth, OAT Growth, Teacher Implementation, SEC, ELLCO, 90 Minute Fidelity. ELLCO was retained in the measurement model as the only measure of teacher practices for which reliabilities could be estimated for this dataset. Two factors emerged, possibly reflecting the dual nature of classroom impact; teacher and student (see Table8).

Table 8

Principal Component Analysis for the Latent Construct of Classroom Impact

Items	1	2
Oat Growth Mean	**0.71**	0.185
Percent Teacher Implement	**0.642**	0.389
SEC	0.628	**-0.423**
ELLCO	-0.279	**-0.605**
Terra Nova Growth Mean	**-0.456**	**0.624**
Eigen Values	1.596	1.119
% of Variance	31.929	22.389

Note: Factor loadings over .4 appears in bold

Classroom Impact demonstrated no statistical significance (Chi square = 4.716, df = 6, p=.581). Other fit indices are CMIN of 4.716, CMIN/DF of .786, CFI of 1.0 and RMSEA of .000 (see Figure 6).
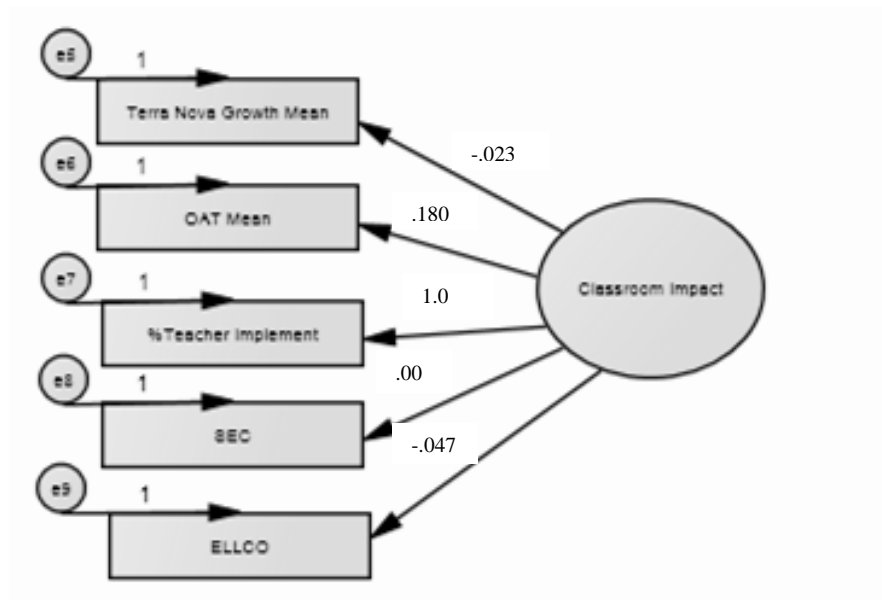
Figure 6.Classroom Measurement Model

Structural Model

While the sample size is small for SEM (n=145), it is within the range of acceptable sample size (Santoro & Bentler, 2001). However, the sample size alone presents the potential for a negatively biased parameter estimation for goodness-of-fit. The measurement models failed to demonstrate a non-significant Chi-square statistic meaning that there is a statistically significant difference between the measurement models and the original dataset. This might be contained within measurement error due to the multiple measures, metrics and time periods in which the measures were collected. However, Jaccard and Wan, 1996 suggest that in the case of small samples (n<200) that additional indicators of fit might provide additional information about model fit. The latent variables and indicators are presented in Table 9.

Table 9

Latent Variables with Indicators and Means Sorted by Measurement Model

Latent	Indicator(s)	Mean
Classroom	TNGMean	1.22
Impact	OAT	400.42
	ELLCO	53.50
	SEC	.1961
	%TIMP	80.13
Leadership	TFullPM	5.9
	%BLDGIMP	85.80
	%PSUPPT	78.24
	FREQCWT	1.74
Ownership	EXTRFO	1.24

Goodness of Fit Indices

In an exploratory manner, and with appropriate caution in any following interpretation of the structural model, the full model was executed. Model trimming was executed to achieve reasonable modification indices (as appropriate given the research and the program) in the face of a significant Chi-square. The final model executed was a recursive model with n=145. The model contained 10 observed, endogenous variables and one unobserved, endogenous variable (Ownership). The model also contained13 unobserved, exogenous variables; Leadership, Classroom Impact and 11 error terms

associated with observed variables. The model contained a total of 65 sample moments with 30 parameters to be estimated. Degrees of freedom were 35. Minimum run was achieved with a significant Chi square of 82.201, df=35, p<.05.

The statistically significant Chi square is indicative of a poor fitting model; however, Chi square is affected by sample size. Researchers suggest that examination of the CMIN/DF, Chi square fit index divided by degrees of freedom, is a more appropriate measure of model fit (see Table 10).

Table 10

CMIN

Model	CMIN	DF	P	CMIN/DF
Default Model	82.201	35	.000	2.349
Saturated Model	.000	0		
Independence Model	420.159	55	.000	7.639

Ullman (2001) states that a value of 2 or less is a good fit while Kline (1998) states that a value of 3 or less is acceptable. The most liberal value accepted is that of 5 or less designated by Schumaker and Lomax, 2004) The default model value of 2.349 is in the acceptable range and is well below the independence model value of 7.639 which is also desirable. (see Table 9) It is important to note that goodness-of-fit indices are affected by sample size also. When a sample is less than 200, goodness-of-fit indices may overestimate in some cases. CMIN, CFI, RMSEA are less affected by sample size than others. It is suggested that at least three indices be reported to adequately address information needs (Jaccard & Wan, 1996). Review of additional fit indices suggested

that the model fit may approach the data set, and the model is logically aligned with program implementation and the literature.

Baseline Comparisons

Baseline Comparisons compare the default model to the null or independence model. The independence model is the worst case in fit with the Chi Square being maximum possible. The comparative fit index (CFI) compares the covariance matrix predicted to the observed covariance matrix while assuming all indicator and latent variables are uncorrelated (see Table 11).

CFI values ranges from 0 to 1 with values approaching 1.0 being a very good fit. By convention, 0.90 is required to accept the model unless there are apriori models in the literature that suggest the value achieved is an improved model. Since there are no existing models in the literature, the default model value of .871 is close but below an acceptable value. The normed fit index (NFI) was developed as an alternative to the CFI. NFI reflects the proportion to which the default model improves upon the null. NFI is reported as a decimal but it represents a percentage so the default value of 0.804 in this study represents an 80.4% improvement over the null or independence model.

Table 11

Baseline Comparisons

Model	NFI	CFI
Default Model	.804	.871
Saturated Model	1.000	1.0
Independence Model	.000	.000

Root Mean Square Error of Approximation (RMSEA) is another check of model fit that is less influenced by sample size than the Chi square. RMSEA corrects for model complexity and penalizes overly complex models (low parsimony). RMSEA values are expected to be in the range of 0.05-0.08, with a good model fit at 0.05 (Schumaker & Loma, 2004). RMSEA is generally reported with a 90% confidence interval and a value indicating the probability of achieving the value in the model (PCLOSE) (see Table 12).

Table 12

Root Mean Square Error of Approximation

Model	RMSEA	LO 90	HI 90	PCLOSE
Default Model	.097	.070	.124	.004
Independence Model	.215	.196	.234	.000

The study RMSEA is 0.097 with a 90% confidence interval of 0.070-0.124 at p=.004. The study model approaches an acceptable figure given that the confidence interval would place the study value within the acceptable range.

Path Estimates

In addition to reviewing fit indices, path estimates are evaluated to determine if the model fit can be improved. Trimming paths can improve model fit and can increase the parsimony of the model. Once a model is fitted (Chi Square = ns, and fit indices within accepted values), then reviewing paths for significant critical ratios (CR) is appropriate. All paths in the measurement models are significant with the exception of the path for SEC to Classroom (CR=1.562), the path of ELLCO to Classroom (CR = -1.248), and FREQCWT to Leadership (CR = -.940) (see Table 13). A CR value below 1.96 indicates that the value is not significant at the 0.05 level. The paths would be areas to consider for

trimming to improve overall model fit. Trimming the non-significant paths in the default

model did not improve model fit. The CMIN/DF increased dramatically when paths were

trimmed (CMIN/DF=7.83).

Table 13

Default Model Path Estimates with Standard Error, Critical Ratio and Probability

Path	Estimate	S.E.	CR	P
Classroom - Ownership	.062	.021	2.895	.004
Leadership – Ownership	-.027	.022	-1.182	.237
%BLDGIMP-Leadership	1.00 (fixed)			
%PSUPPT	.693	.099	7.026	***
TFullPM – Leadership	.114	.016	7.166	***
EXTRFO – Ownership	1.00 (fixed)			
FREQCWT – Leadership	.080	.018	4.587	***
TNGMEAN-Classroom	-.023	.025	-.940	.347
%TIMP - Classroom	1.00 (fixed)			
ELLCO - Classroom	-.047	.037	-1.248	.212
OAT – Classroom	.180	.046	3.918	***
SEC – Classroom	.000	.000	1.562	.118

*** *indicates significant at less than .001*

In addition to the measurement models, the structural model paths are estimated.

The path from Leadership to Ownership was not significant (CR = -1.182). In an effort

to improve the fit, the covariance between the two latent constructs was fixed to 1

(standardized). This resulted in improved critical ratio (CR) for the path from Leadership

to Ownership but negatively impacted the overall fit of the model on the fit indices. In the

final model, the covariance path between the two latent constructs was allowed to freely

vary as was necessary to achieve a more closely fitting model (see Figure 7).
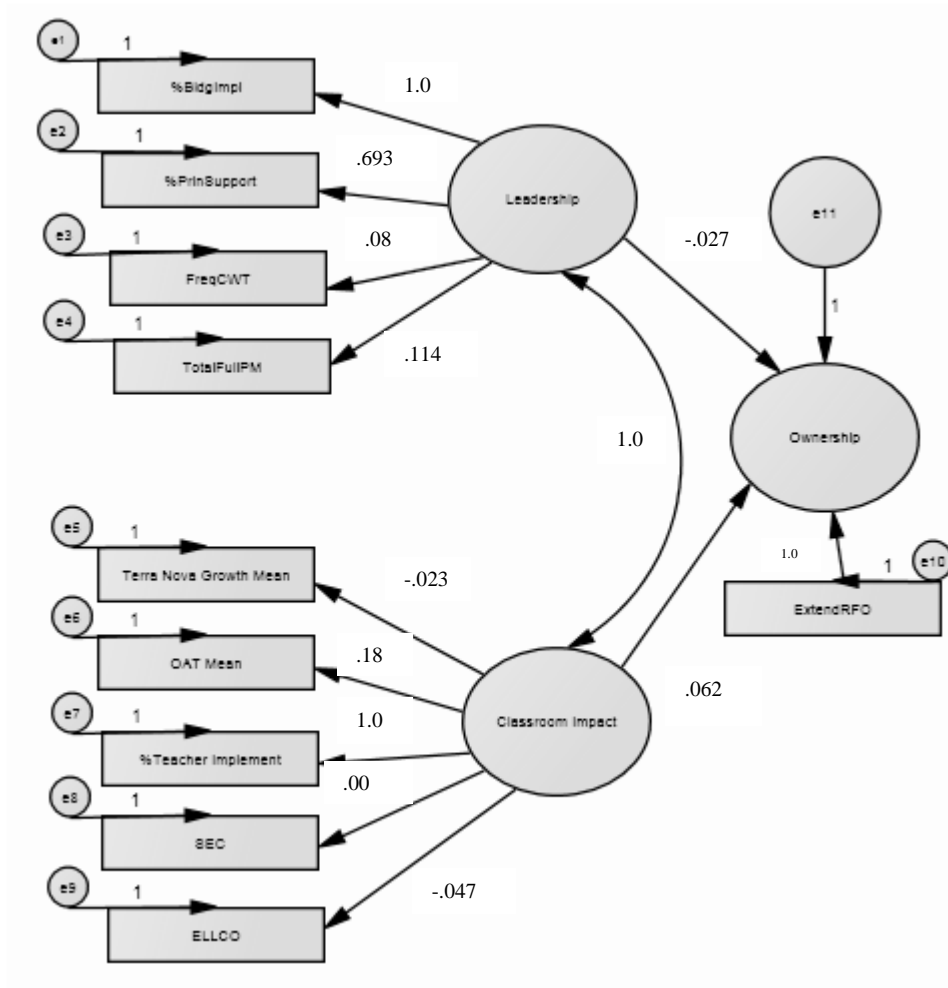
Figure 7.Final Structural Model

General Research Questions

To what extent does the theory-based sustainability model accurately

reflect the reform experience of the identified Reading First Ohio schools?

The archival nature of the data caused some adjustments to the model. The SEM

model created by this study is a poorly fitting model. With a significant Chi square of

82.201, df=35, p<.05. Per Jaccard and Wan (1996), Goodness of Fit indices were

reviewed in the face of the overall statistical significance of the model. According to

Kline (1998) the CMIN/DF is examined. The value achieved was 2.349 which is an

acceptable value (Kline, 1998). The CFI value was .871, which approaches the

acceptable value of .90. The NFI value was .804, demonstrating that the model produces

was 80.4% improved over the null or independent model. Ninety percent or .90 value on

the NFI is the generally accepted standard for reasonable fit. Path estimates

demonstrated a CR of over 1.96 (2.895) for the Classroom to Ownership path but not for

Leadership to Ownership. There are multiple possible explanations for the weak model

fit that will be discussed in Chapter V.

> *To what extent does the model accurately predict the districts most likely*
>
> *to sustain their reform effort as identified by the degree of ownership*
>
> *demonstrated?*

To test this question, a regression equation was created using the beta weights

estimated by the SEM from this study. The predicted values were then used to rank order

the districts. Kendall Tau was used to test the correlation of the results.

> *To what extent does is the model able to accurately predict the*
>
> *performance of the schools relative to expert judgment?*

Regional Consultant is the title given to individuals who served as technical

assistance and implementation accountability personnel for Reading First Ohio. Seven

individuals were trained in program monitoring using the Program Monitoring Tool

(Salzman & Newman, 2003) developed specifically for the purpose of use of Reading

First. Each of the Regional Consultants were assigned to multiple districts, multiple

elementary buildings. Each Reading First Ohio Regional Consultant has participated in

Reading First Ohio for a minimum of four years. Each participant elementary building was monitored three times per year for each year in the grant at full funding, and each was monitored twice per year during the final or sustainability planning year. Inter-rater reliability was estimated at 0.85.

As part of the ongoing program implementation and efforts to provide technical assistance, Regional Consultants were asked to create a projection of the expectation for district sustainable program efforts. The goal of the process was to prepare to customize sustainability conversations for the districts. As part of the process, the Regional Consultants were asked to place districts on a scale of implementation that progressed beyond basic program monitoring levels into exploration of higher degrees of ownership. The districts were placed on a continuum of 1-10: 1= minimal program implementation-problematic level, 4- basic levels of program implementation, 5-7 progressively more authenticity in implementation, and above 7 demonstrating increasing levels of ownership and adaptation of the program to unique district contexts.   Based on their professional perspective on the program implementation and on the outcomes of the Program Monitoring, each Regional Consultant rated their own districts. The Regional Consultants then engaged in a question and answer process further refining the meaning of the scale and of the technical assistance needs in order to achieve a sustainable program. The differences and similarities between the two rankings is apparent when they are viewed together (see Table 14).

Table 14

Comparison of Ranking of Districts Likely to Sustain Using the Ownership Value

Predicted by the Model and Using Expert Ranking Based on Program Monitoring

Activities

District	Predicted Ownership Value	RC Ranking	RC Value
Southern Local	2.694	Vinton County	10.0
Belpre	2.683	New Boston	10.0
New Lexington	2.655	Crooksville	10.0
New Boston	2.650	Belpre	8.0
Alexander	2.637	Southern	8.0
Crooksville	2.620	Alexander	8.0
Lockland	2.612	Portsmouth	8.0
Sandusky	2.605	New Lexington	8.0
Federal Hocking	2.594	Jefferson	8.0
Vinton County	2.578	Sandusky	8.0
Middletown	2.573	Lockland	7.5
Portsmouth	2.536	Conneaut	7.0
Jefferson	2.539	Columbus	6.5
Marion	2.532	Federal Hocking	6.0

| | Predicted Ownership | | |
District	Value	RC Ranking	RC Value
Frontier	2.300	Marion	5.5
Toledo	2.286	Dayton	5.5
East Cleveland	2.224	Toledo	5.5
Dayton	1.941	Springfield	5.0

Reviewing the relative rankings of the districts by quartile the degree of agreement between the two can be identified. Quartile 1 demonstrated a 60% agreement between the two rankings, Quartile 2 demonstrated 20% agreement, Quartile 3 demonstrated 0% agreement, and Quartile 4 demonstrated 60% agreement. If taken by halves, the higher performing half demonstrated 80% agreement, and the lower performing half demonstrated 70% agreement. The "Ownership" score generated using the model was compared with the scoring of the Regional Consultants (See Table 14). There is a high degree of alignment. A post-hoc analysis using Kendall's tau non-parametric correlation was conducted to examine the relationship between the scoring on "Ownership" predicted by the model and the expert rating for sustainability. There was a statistically significant correlation between the two ratings. ($\tau=.489$, n=20, p<.01)

To what extent do the latent variables of ownership and collaborative adaptation impact the overall functioning of the model?

Due to the constraints imposed by the database, collaborative adaptation was eliminated from the study. Given the general failure of the model it is impossible to adequately address this question with the current study.

To what degree do the identified challenges of leadership mobility and construction impact the ability of the school/district to achieve a fully implemented and sustainable model?

During the conceptualization of this study, it was noted that a number of the participating Reading First sites were experiencing one or more challenges related to leadership change and/or construction projects on site. Leadership changes took the form of both superintendents and building principals. During the course of the program implementation the state of Ohio initiated a significant number of construction projects within the schools of the state. The Reading First Ohio schools were impacted by the construction. Construction took the form of new buildings often initiating redistricting, complete renovation of existing buildings that required relocation or major reconstruction that took place in occupied schools. Data collected demonstrated the degree to which leadership was mobile during the course of Reading First. The principals changed at least once in 62.41% of the elementary buildings with several schools experiencing multiple principal turnovers. One school had 4 different principals during the six-year initiative. District leadership was also mobile, with 54.89% of the districts reporting superintendent changes during the initiative. Construction was another form of impact on the initiative. Construction impacted 56.6% of the elementary schools during the initiative. Construction varied from renovation occurring within the building and requiring modifications to routines, all the way to building closings requiring redistribution of students and personnel.

Although these variables did not load into the SEM model an examination of the correlation between student outcomes and construction, principal change and

superintendent change provides an understanding of the impact of these issues (see Table 15).

Table 14

Correlation of Student Achievement Outcomes

	OAT	
Variable	Mean	Sig.
% P Change	-.237**	.002
% Construction	-.282**	.003
%SuptChange	-.002	.984

Summary

Chapter IV provides an overview of the analysis conducted for this study. Due to limitations in the archival data it was necessary to alter the intended theoretical model to be studied. Adaptive Collaboration was not sufficiently documented in the dataset to retain it as a construct in the model. The second section provides details of the model development process. A principal components factor analysis was conducted to illuminate the underlying factor structure prior to creation of the measurement model. Initial factors structure indicated that the strongest factors might be only two latent constructs; Leadership (and Program Fidelity as one construct) and Classroom Impact. Classroom impact appeared to reflect the dual areas of classroom change; teacher and student. The measurement models proved difficult to fit. Leadership with four observed variables proved to be an inadequate fit. Classroom Impact demonstrated satisfactory properties.

The structural equation model was evaluated with caution given the failure of the measurement model; Leadership construct. The structural model was statistically significant with a Chi Square = 82.201, df=35, p< .05. Other goodness-of-fit indices produced a model with closer approximation of the underlying dataset with a CMIN/DF of 2.39, a CFI of .871 and an RMSEA of .097. NFI demonstrated that the default model improved upon the null by 80.4%. Path estimates demonstrated a CR of over 1.96 (2.895) for the Classroom to Ownership path but not for Leadership to Ownership. Discussion of the model will be continued in Chapter V.

CHAPTER V

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

This chapter provides a brief summary restating the problem and purpose of the study, an overview of the methodology and hypotheses, conclusions and discussion of the findings, followed by implications, limitations, and concludes with recommendations for further research.

Summary of the Study

The *American Recovery and Reinvestment Act of 2009* allocated $5 billion to early education and another $77 billion to elementary and secondary education. Within that budget is $48.6 billion to stabilize our state education budgets and another 5 billion dollars to foster and encourage innovative education practices. In 2010 the U.S. Department of Education allocated $4.35 billion to the competitive grant allocation for *Race To The Top*. In May 2010, they allocated $650 million to *Investing in Innovation (i3)* grants. Each of the funding cycles lasts from 1-5 years. Each requires an application and a plan for implementation, documentation and accountability. Applicants are asked to consider the issue of sustainability but specific plans to sustain are low on the scoring criteria (U.S. Dept. of Education, 2010). What do we expect as a long term return on

investment in education?   What are the lasting outcomes of innovation, reform and renewal in education?

Farrell (2003) described government based funding as the primary catalyst for school change in many educational settings adding that this process of cyclic efforts has created a revolving door of school reform efforts that align only with funding cycles. Fullan (2001) challenged funding driven reform and discussed innovation as the appropriate catalyst for change. He described the many pressures applied in the initiation of an innovation process that often do not promote full implementation let alone long-term change. Educational organizations are inundated with reform initiatives that are launched and abandoned with regularity. He stated that seldom are programs planned and implemented with deliberate intent to sustain and continue. Generally, the focus for schools is the implementation of model initiatives, not on long term or lasting impact. Successful outcomes and full program implementation are necessary but not sufficient indicators of lasting change (Datnow, 2005).Lasting change, *sustainability*, has remained an elusive target even in the face of initially successful implementation (Datnow & Stringfield, 2000).

This study reviewed the literature on school change and the factors that most impact that change. Teachers are a critical factor in school change. They work at the point where theory and practice come together. Further, they are joined at that juncture by a group of young people who bring with them their own capacities.   Successful school initiatives consistently demonstrate that teachers must be highly knowledgeable in their content and have time for professional collaboration. The classroom is the basic building block to any successful change program. (Darling-Hammond, 2010) Teacher certainty, self esteem,

and isolation impact their ability to implement change processes in the classroom (Rosenholtz, 1989, Kruse, Louis & Bryk, 1995; Newmann & Wehlage, 1995).

Multiple researchers have identified the vital role that leadership plays in executing and sustaining school reform. Building leadership has been positively linked to student outcomes (Nettles & Herrington, 2007), to workplace relationships (Whitaker, 2004), and to creating an adaptive organization (Fullan,2001; Heifetz & Linsky, 2002) Beyond leadership in the school building is district leadership, central office as the policy body can either support or undermine sustainable school reform. Policies that chase funding and that encourage multiple change efforts in rapid succession create a cycle of reform efforts that cannot be fully implemented or sustained at the building level (Little & Dorph, 1998, Spillane, 2000). Spillane (2000) further stated that even when the district endorsed and supported the reform efforts, lack of knowledge or understanding on the part of policy makers often contributed to weak, piecemeal and ineffective programs.

Sustained school reform efforts are few but even when a program is sustained, there is limited opportunity to learn from it. Funding cycles, and the associated program evaluation, generally end with the pre-determined implementation cycle leaving few resources for the monitoring and data collection of a long term sustainability study. Studies of sustained school change require long-term research efforts. Longitudinal studies, traditionally limited by time, resources, and subject availability, are less frequently conducted than other research designs (Bijleveld & Leo, 1998). As a result, sustainable school reform is an area of concern for many but researched by few. There are research studies and there are program evaluations. Program evaluations focus on practice, research studies build theory. Hatch and White (2002) stated that the body of

literature surrounding school improvement is too fragmented to provide the basis for academic discussion and clear understanding of school reform issues. "Despite widespread conviction that adequate knowledge exists for improving schools, we argue that the knowledge needed for successful school reform goes far beyond what is currently available and accessible" (p. 117).

The purpose of this study was to build to the body of literature around the issue of sustainable school change. In this investigation, program implementation and successful outcomes were incorporated as variables in the larger issue of sustainable change efforts. Reading First Ohio data, years 2003-2009, were used to build a theory-based model of sustainable school change demonstrating that a prerequisite to sustainable efforts is the ownership of the change process by school personnel and to test the hypothesized model against the progress of Reading First Ohio schools toward sustained efforts. Providing empirical evidence of the contributory nature of ownership will allow this study to build on sustainability in school change literature.

Summary of Findings

This section is organized by general and specific research hypotheses. In each category, conclusions and discussion will be presented for those hypotheses.

Hypothesis 1.

The proposed theoretic model was determined to be statistically significantly different from the underlying dataset. (Chi Square = 82.201, df =35, p<.05) Review of additional fit indices as appropriate according to Jaccard and Wan (1996), demonstrated that the model achieved an adequate CMIN/DF (2.349), CFI (.871) and approached an adequate RMSEA (.097). According to NFI, the null model was improved upon by

106

80.4%, although 90% is the generally accepted standard for a good fit on a model. Further modifications to improve the fit were neither fruitful nor appropriate. SEM is one of the few statistical applications where the results are likely to show that the original theoretic model is in error. Once the model respecification process fails to improve the model, the researcher needs to stop modifying. "If it is true that a proposed model does not reflect reality, then reaching a conclusion of misfit between data and model should be a desirable goal, not one to be avoided by careless respecifications until satisfactory levels of fit are achieved" (Mueller & Hancock, 2008, p 506).

While it would be inappropriate to interpret the model for causal effects, cautious interpretation of the results may be warranted given that the fit indices met or approached minimum criteria. SEM requires a solid balance between the analysis conducted and the underlying body of literature from which it emerges. Given the failure to reach a good fit on the model, caution should be exercised in the interpretation of the results. However, theories and models that fail contribute to the body of knowledge; thus, interpretation of the failure of the model and the possible underlying causes of the model failure could provide insights.

Logically, and in alignment with the Reading First Ohio initiative design, the results of the study provide an opportunity for discussion of the theoretic model. The program design for Reading First Ohio emphasized impacts on the classroom. The significance (CR=2.895, P=.004) of the path between the construct "Classroom" and the "Ownership" is both logical and appropriate given the design of the initiative and the implementation. The lack of significance in the path between the construct "Leadership" and "Ownership" (CR= -1.182, p<.237) may be indicative of program design issues as

well. The program design was limited in its involvement with building leadership. In addition to program design questions, there were also limited data collected about the leadership development that existed. There were monthly minutes of professional development mandated for teachers but there were more limited expectations for professional development for principals. Program monitoring tracked closely the professional development for teachers and the implementation of the initiative in the classroom. Program monitoring did not monitor principal involvement beyond basic attendance at meetings and communications. Classroom walkthrough training was provided to principals, but implementation was not a grant requirement so compliance was inconsistent. As a routine of the grant large amounts of student data were collected, less but still substantial amounts of data concerning teachers were collected but little data were collected pertaining to building leadership. Further, the highest level of missing data, even when collected, was in the leadership area. The performance of the measurement models, structural model and path values may reflect this discrepancy in the structure of the program.

Although a generally poor fitting model, there is reason to believe that portions of the model may accurately reflect at least portions of the experience of Reading First and may also highlight some of the data collection and programmatic concerns.

Hypothesis 2.

In an exploratory manner, the regression weights for all variables contributing to the model at a statistically significant level were used to predict district outcomes on Ownership. The ownership values produced were consistent with Reading First Ohio district monitoring and anecdotal experiences of Reading First personnel. It is also

noteworthy that of the 10 highest "ownership" values predicted, 8 of them occurred in small districts with the ODE designation of "Rural Agricultural, high poverty, low median income". Several factors might contributed to this; teachers in these schools and districts are likely to live and work in the same communities as their students, the identified districts have traditionally had few opportunities to participate in large federal initiatives offered in Ohio and therefore "geared up" at a higher level, and/or these schools and districts tend to be smaller and more personal thus generating a greater sense of ownership of the school. This is an area for further exploration as it may have implications for community engagement and community organizing around school reform activities.

Hypothesis 3.

As part of the technical assistance planning process the Regional Consultants were asked to rate the districts that they monitored and provided with technical assistance. They were asked to rate the districts on a scale that was collaboratively developed as part of technical assistance planning for sustainability of Reading First following the close of the grant. The scale was a 1-10 continuum. A score of 1 indicates a non-compliant or new- start program and a high score of 10 indicates those districts that are ready to move to sustainability. This ranking was later used as the basis of the comparison against the ranking generated by the model ownership scores.

The "Ownership" score generated using the model was compared with the scoring of the Regional Consultants (see Table 14). There is a high degree of alignment. A post-hoc analysis using Kendall's tau non-parametric correlation was conducted to examine the relationship between the scoring on "Ownership" predicted by the model and the

expert rating for sustainability. There was a statistically significant correlation between the two ratings. (τ=.489, n=20, p<.01)

Although the SEM is weak, it appears that the predicted values on ownership were aligned with the scoring for sustainability that were completed by the Regional Consultants as expert judges thereby demonstrating some concurrent validity.

Hypothesis 4.

This hypothesis could not be investigated due to the lack of data for the variable of collaborative adaptation. Since this was intended to have one observed variable associated with it in the model, there was no way to scale the variable. The literature supports the need for ownership to create sustained change. Ownership is the process of shifting from an externally driven program to one that is internally owned and directed. Colburn (2008) stressed the need to link ownership and sustainability in school reform efforts and emphasized that one is unlikely without the other. The lack of ability to test this area leaves this as an open area for future research.

Hypothesis 5.

Barker (2006) studied the impact of leadership mobility on reform efforts and, as expected, demonstrated that stable leadership plays a central role in positive outcomes. Changing leaders diminished an organization's ability to move a reform process forward and to sustain organizational culture. A study by the Rand Corporation (1977) reviewed 293 federally funded projects and found that principal and staff turnover was one of the most significant factors in abandoning new reform efforts. In the Reading First initiative, 54.89% of the districts experienced a superintendent change, 62.41% of the buildings experienced a leadership change and 56.6% of the buildings experienced construction.

Construction projects ranged from renovations that required changes to school routines to complete new facilities that required students and personnel to relocate. Construction also caused redistricting in several districts.

Although the mobility of leaders and the construction did not enter into the model, a review of the correlation between student outcomes (OAT scores) and the building leadership change (r= -.237, p=.002) showed that principal mobility was negatively correlated with student outcomes in Reading First Ohio. A review of the correlation between %Construction and student outcomes also demonstrated a significant negative relationship (r= -282, p=.003). The relationship between superintendent change and student outcomes was negative but not significant. (r= -.002, p=.984)

Given the strength of the correlations, the degree to which Reading First districts maintained the school change effort and were able to move toward ownership is a positive. This further strengthens the model stating that when classroom impact and teacher buy-in is high, the negative impact of leadership change may be diminished.

Limitations

The first limitation of the study was the dataset selected for analysis. Like most school change initiatives, Reading First Ohio did not build the program with specific plans for sustainability. Sustainability discussions were brought into the program mid-implementation and so data specifically related to sustainability were not specifically structured into the data collection as part of the design. This meant that selection of variables had to be back-fitted to meet the needs of this study.

Archival databases have inherent concerns. The dataset was collected and useful for Reading First Ohio. The database was collected by multiple agencies under a variety

111

of conditions. The operational definitions, instruments and data collection processes served the purpose for which it was designed but did not adapt easily to the needs of this study. For some measures, the standard error of measurement was high due in part to the multiple purposes and agencies involved in the data collection process. The archival nature of the data also limited the researcher's ability to limit missing data or to seek to improve the fidelity of the collection process. Missing data, both within the data collected and by what remained uncollected, were a problem. The sample size itself may also be a limitation to the study.

SEM as an analysis tool has inherent limitations. SEM provides an opportunity to review the underlying data structures and supplement theory, but there is also the risk of moving away from the original theory base as the structural equation model is modified to achieve a closer fit. Spanos (1986) cautioned researchers against the temptation of moving the model too far away from the theoretic and literature based structure to that of the data generated by the study. The process of model modification must be guided by the theory first and by the SEM model fit second. Moving too far from the original theory base threatens to create an illogical model, regardless of the statistical fit. SEM is one of the few statistical applications where the results are likely to show that the original theoretic model is in error. "If it is true that a proposed model does not reflect reality, then reaching a conclusion of misfit between data and model should be a desirable goal, not one to be avoided by careless respecifications until satisfactory levels of fit are achieved" (Mueller & Hancock, 2008, p 506).

An additional limitation is that of the research on school reform. In addition to the limitations of funding and resources mentioned earlier, there are additional design

concerns including lack of control groups, problems associated with multi-site implementation and data collection (Pogrow, 1998; 2000). Sample size is also a concern, as seen in this study. Studies that focus on the school or district as the unit of analysis suffer from low sample sizes compared to studying students or teachers as the unit of analysis. Finally, school change is a dynamic process that is adapted as a process moves forward. This means that strict research controls and definitions are often counter-productive and nearly impossible (Cuban, 1993, Datnow & Stringfield, 2000; Fullan, 2000).

Finally, literature on school reform has focused primarily on testing reform models and demonstrating those factors that impact outcomes. There is little literature currently available to guide the development of sustainability theory or models. This is an area for development in the literature.

Implications

Research.

At this time, school reform, school change and new initiatives are omnipresent in education. Billions of dollars are invested each year in efforts to produce change, generate innovation or increase "best practices" in education. Researchers need to begin to identify key variables that contribute to sustainability and to develop models that address long term sustainability of educational initiatives. Researchers need to seek examples of sustained change and examine the characteristics of the innovation, of the organization and of the resources that supported that change.

School reform efforts.

There are implications from this study for school reform planning and implementation. The model did not demonstrate full efficiency as a model of sustainable change but there are still implications for program planning. The significant paths reinforced the school change literature and confirmed the need for school change models to pay particular attention to changes in the classroom. Change in the classroom was a clear path to ownership of the change process. Addressing the learning needs of teachers and students, creating a positive climate and culture and creating an environment of trust are useful investments for school reform models. Further the model suggested, although the data were not sufficiently explained, that there is a Leadership component to be addressed in school reform efforts. Like Reading First, too many initiatives address the issue of leadership as an addendum not as a central element of a change initiative. The failure to demonstrate the impact of Leadership on ownership and therefore on sustainability may have been negatively impacted by several factors such as the program design, the data collection or by a true lack of impact. Since the Leadership paths could not be trimmed from the model completely, and based on the literature, it is reasonable to assume that these paths are an important area for continued research. In school reform research we are equally unsure about causes of failures as we are about successes and sustainability. In the failure of this model, there are questions raised that will promote additional research into the process of change.

Policy makers.

Finally, there are implications for policy makers. School reform agendas need to shift focus from assuming change will last to planning for lasting change (Fullan,

1999).The current cyclic funding does not contribute to long-term improved educational outcomes. Research into change models can provide foundations upon which educational innovations can be built. Program initiatives can be designed for the greatest likelihood of long term impact. Policy makers need to include sustainability planning into funding requests in a purposeful and meaningful manner. In addition, there needs to be a shift in funding and monitoring that will allow planning for sustainability and follow-up on long term efforts.

Recommendations for Further Research

This study has the opportunity to lay the foundation for additional research into school reform sustainability and the role of ownership in creating that sustainability. The need for researchers to develop and test models of sustainability is a broad area of need.

One possibility for future research with this particular database would be to replicate the current study as closely as possible using multiple sites as separate samples and adjusting the model to reflect the shift. The limitations of this data set would have to be addressed and some additional data collected, which might be possible with fewer selected sites.   In addition, replicated in this way, the researcher could elect to identify the participant sites by specific characteristics such as type of district (rural, urban, major urban) or the rank scoring on ownership.

An area that was not directly addressed in this study was that of community and school change sustainability. The emergence of the district type as a possible factor in the degree of ownership displayed by Reading First districts may be indicative of a community and school interaction. Districts where the school personnel live and work in the same community as the school district may have a higher degree of ownership prior

115

to any school change efforts. District personnel interact with community members on a regular basis through daily activities such as shopping, recreation and social activities. The shared sense of community and increased ownership in school outcomes is often demonstrated in the manner in which school personnel discuss the students; "our kids" vs. "the students" and "our school" vs. "the school." Even in the poorest communities a trusting relationship between school and community increases the likelihood that any school improvement process will be successful. Based on the experience of Reading First Ohio, an additional interaction in this area might be cultural. Rural south-east Ohio is Appalachian and has some of the underlying cultural norms and values that extend from that heritage. Those cultural differences may contribute to the increased ownership demonstrated by the Appalachian districts participating in the initiative.

Testing of the theoretic model with entirely different data from an unrelated reform initiative might also provide insight into the functioning of the model and the impact of the construct of ownership. This might be accomplished in two different ways:1) test against an existing initiative, 2) test against an initiative in which the data has already been archived (as in a government database). The archived data study could then include a limited data collection to determine the degree to which the program or initiative had been sustained.

The issue of collaboration and collaborative actions to adapt a reform effort to the needs of the organization requires further attention in the research literature. Collaboration is widely acknowledged as necessary in school reform work (Darling-Hammond, 2010) but there is little empirical research documenting the specific contribution of collaboration to the school reform process. Top down methods of school

reform failed to yield the types of improvements desired (Datnow, 2000) while collaborative processes are seen as a more productive method for advancing an initiative (Timperley & Parr, 2005). The Reading First schools that demonstrated higher levels of collaboration within the teaching and professional staff were also able to demonstrate higher levels of success in classroom change and school improvement

Finally, additional work needs to be done to develop the concept of ownership and the relationship between ownership and sustainability. That research needs to be directed to the study of the issue of ownership and change processes within given contexts and specific types of actions or initiatives. Further, another area for study is that of the relationship between ownership and organizational capacity for relatively small incremental changes that is: adaptation. This has implications not only for school change but for organizational and community change initiatives as well.

Summary

This study investigated the role of ownership in sustained school change using the experiences of the Reading First Ohio initiative. The dataset used to test the theoretic model was collected as part of the Reading First Ohio initiative between the years 2003-2009 and was a composite of data collected by the Reading First initiative and the Ohio Department of Education. The unit of analysis for the study was the 145 elementary school sites (IRN determined).

A two stage model development process was utilized: measurement model and structural model. The Principal Components factor analysis provided insight into underlying factor structure which reinforced the theoretic model in some areas. Based on factor loadings two measurement models were created: Classroom Impact and

117

Leadership. When tested the measurement model for Classroom Impact was not significant, the measurement model for Leadership was. Respecification of the Leadership measurement model failed to improve the fit. In an exploratory manner and with appropriate caution in interpretation, the structural model was tested. The model was statistically significant demonstrating a difference between the theoretic model and the underlying dataset. Researchers suggest that even in the case of a significant chi square that review of addition fit indices might provide additional information about the model functioning. (Jaccuard & Wan, 1996) In the case of this theoretic model the additional indices demonstrated some elements where fit achieved or approached minimum standards. Further, when tested against the experience of the Reading First initiative there were areas that demonstrated alignment of the model and experience.

Although weak, the model achieved acceptable scores on several fit indices and was parsimonious as well. It was able to demonstrate some concurrent validity with expert judgment and supported the overall design and implementation of the underlying program model. The results mimic the experience of Reading First and suggest that further research in the areas of sustainability and ownership are needed.

There are policy implications for developing models of school reform sustainability that could inform future funding initiatives, future program design and implementation as well as technical assistance in support of the initiatives. Ultimately, planning for sustainability will impact the long term outcomes of educational programs.

REFERENCES

Anderson, J.C. & Gerbing, D.W. (1988). Structural equation modeling in practice: A review and recommended two step approach. *Psychological Bulletin, 103*, 411-423.

Arbuckle, J. L. (2007). *AMOS*. Chicago, IL: SPSS.

Barker, B. (2006). Rethinking leadership and change: a case study in leadership succession and its impact on school transformation. *Cambridge Journal of Education, 36*(2), 277-293.

Bentler, P. M. (2006). *EQS*. Encino, CA: Multivariate software.

Bentler, P.M. & Chou, C.P. (1987). Practical issues in structural equation modeling. *Sociological Methods & Research, 16*, 78-117.

Berman, P., & McLaughlin, M. (1977).*Federal programs supporting educational change, vol 7: Factors affecting implementation and continuation*. Santa Monica, CA: Rand Corporation.

Bijeveld, C.J.H., & Leo, J. (1998).*Longitudinal data analysis: designs, models and methods*. Thousand Oaks, CA: Sage.

Blank, R.K. (2002).Using Surveys of Enacted Curriculum to advance evaluation of instruction in relation to standards. *Peabody Journal of Education, 77*(4) 86-121.

Bodilly, S. J., Schuyler, G. Reichardt, R. (1998). *Lessons from new American schools' scale-up phase: prospects for bringing designs to multiple schools*. Washington DC: Rand.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley & Sons.

Bryk, A., Sebring, P., Kerbow, D., Rollow, S., & Easton, J. (1998). *Charting Chicago school reform*. Boulder, CO: Westview Press.

Byrne, B.M. (2010). *Structural equation modeling with AMOS*. New York: Taylor and Francis Group, LLC.

Colburn, C.E. (2008). *Supporting and sustaining Reading First: the roll of central office*. Sustainability meeting, September 2008. Columbus Ohio: 23 pages PPT.

Colburn, C.E. (2004). Beyond relationships between the institutional environment and the classroom. *Sociology of Education, 77*, 211-244

Colburn, C.E. (2003). Rethinking scale: Moving beyond numbers to deep, lasting change. *Educational Researcher, 32*(6), 3-8.

Darling-Hammond, L. (1997). School reform at the cross-roads: confronting the central issues of teaching. *Educational Policy, 11*(2), 151-66.

Darling-Hammond, L. (2010). *The flat world and education: How America's commitment to equity will determine our future*. New York, NY. Teacher's College Press.

Datnow, A. (2005). The sustainability of comprehensive school reform models in changing district and state contexts. *Educational Administration Quarterly, 41*(1), 121-153.

Datnow, A. (2000). Implementing an externally developed school restructuring design. *Teaching and Change, 7*(2), 147-171.

Datnow, A, Borman, G, & Stringfield, S. (2000). School reform through a high specified curriculum: Implementation and effects of the core knowledge sequence. *The Elementary School Journal, 101*(2), 167-191.

Datnow, A, & Stringfield, S. (2000).Working together for reliable school reform. *Journal of Education for Students Placed at Risk, 5*(1&2), 183-204.

DeBray, E.H. (2006).*Politics, ideology and education*. New York, NY: Teacher College, Columbia University.

Desimone, L. (1992). How can comprehensive school reform models be successfully implemented? *Review of Educational Research, 72*(3), 433-79.

Education Trust. (1999). *Dispelling the myth: high poverty school exceeding expectations*. Policy Paper, Washington, DC.

Elmore, R. (1995). Getting to scale with good educational practice. *Harvard Educational Review, 66*(1), 1-26.

Finn, C., & Hess, F. (2007, Spring). What Innovators can and cannot do. *Education Next*, 1-18.

Fordham Foundation, (2009). *Seeking quality in the face of adversity*. Fordham sponsorship accountability report. Retrieved (2009, September 10) from http://www.edexcellence.net

Friedman, T. (2005).*The world is flat*. New York: Farrar, Strauss and Giroux.

Fullan, M.(2006). The Future of educational change: system thinkers in action. *Journal of Educational Change, 7*, 113-122.

Fullan, M. (2001).*The new meaning of educational change*. New York: Teachers College Press.

Fullan, M. (2000).The return of large scale reform. *The Journal of Educational Change, 1*(1), 1-23.

Fullan, M. (1999).*Change forces: The sequel*. Philadelphia, PA: Falmer Press/Taylor

    Francis, Inc.

Fullan, M.G. (1993, March). Why Teachers must become change agents. *Educational*

    Leadership, 50(6).

Fullan, M. & Hargreaves, A. (1992).*What's worth fighting for? Working together for*

    your school. New York: Teachers College Press.

Gladwell, M. (2000).*The tipping point*. Boston, MA: Little, Brown and Company.

Good, R. H., & Kaminski, R. A. (2002). *DIBELS oral reading fluency passages for first*

    through third grades (Technical Report No. 10). Eugene, OR: University of

    Oregon.

Good, R. H., & Kaminski, R. A. (2002). *Dynamic Indicators of Basic Early Literacy*

    Skills. Eugene, OR: Institute for the Development of Educational Achievement.

Good, R. H., Wallin, J. U., Simmons, D. C., Kame'enui, E. J., & Kaminski, R. A. (2002).

    System-wide percentile ranks for DIBELS benchmark assessment (Technical

    Report No. 9). Eugene, OR: University of Oregon.

Guskey, T., & Huberman, M. (1995).*Professional development in education: new*

    paradigms and practices. New York: Teachers College Press.

Hall, G.E., & Hord, S.M. (1987).*Implementing change: patterns, principals and potholes.*

    New York, NY: Allyn & Bacon.

Hatch, T. (2002, April). What happens when improvement programs collide? P*hi Delta*

    Kappan, 83(8), 626-634.

Hatch, T. (2001). It takes capacity to build capacity. *Education Week,* 12-13

Hatch, T. (2006). Improving schools in turbulent times. *The New Educator, 2*, 267-276.

Hatch, T, & White, N. (2002). The raw materials of reform: rethinking the knowledge of school improvement. *Journal of Educational Change, 3*, 117-134.

Heath, C, & Heath, D. (2007).*Made to stick*. New York: Random House.

Heifetz, R.A., & Linsky, M. (2002).*Leadership on the line: staying alive through the dangers of leading*. Boston, MA: Harvard Business School Publishing.

Hill, P., & Celio, M. (1998).*Fixing urban schools*. Washington, DC: Brookings Institute.

Hooper, D., Coughlan, J. & Mullen, M. (2008) Structural equation modeling: guidelines for determining model fit. *The Electronic Journal of Business Research Methods. 6*,53-60.

Huberman, M., & Miles, M. (1984).*Innovation up close*. New York, NY: Plenum.

Joreskog, K.G, & Sorbom, D. (1988).*LISREL 7: A guide to the program and applications*. Chicago, IL: SPSS, Inc.

Kazdin, A.E. (1995). Preparing and evaluating research reports. *Psychological Assessments, 7*(3), 228-237.

Keltner, B. (1998). *Funding comprehensive school reform*. Washington DC: Rand Corporation.

Kerlinger, F.N., & Lee, H. (2000).F*oundations of behavioral research*. New York: Harcourt College Publishing.

Kruse, S.D., & Louis, K.S. (1993).*Developing professional community in new and restructuring urban schools*. Eric. Retrieved from http://www.eric.ed.gov

LeVine, A.E. (2009, Feb.). Waiting for transformation. *Education Week*, 34-35.

Little, J., & Dorph, R. (1998).*Lessons about comprehensive school reform*. Berkley, CA:University of California Press.

McChesney, J., & Hertling, E. (2000).The Path to comprehensive school reform.

    Educational Leadership, 57(7), 10-15.

McGuinn, P.J. (2006). *No child left behind and the transformation of federal education*

    policy: 1965-2005. Lawrence, KS: University of Kansas Press.

McLaughlin, M., & Talbert, J. (2001).*Professional communities and the work of high*

    school teaching. Chicago, IL: University of Chicago Press.

Mussoline, L. (2001) School restructuring as a policy agenda: why one size may not fit

    all. *Sociology of Education, 74* (January), 724-736.

Mueller, R.O., (1997) Structural equation modeling: back to basics. *Structural Equation*

    Modeling: a multi-disciplinary journal. 4, 353-369.

Mueller, R. O. & Hancock, G. R. (2008). Best practices in structural equation modeling.

    In J. W. Osborne (Ed.), *Best practices in quantitative methods* (pp. 488-508).

    Thousand Oaks, CA: Sage Publications, Inc

Muthen, R.O. & Muthen, L.K. (2006).*Mplus*.Los Angeles, CA: Computer software,.

National Commission on Excellence in Education, NCEE.(1983). A *nation at risk*.

    Washington DC: United States Department of Education.

National Reading Panel. (2000). *Teaching children to read: An evidence-based*

    assessment of the scientific research literature on reading and its implications for

    reading instruction. Bethesda, MD: National Institute of Child Health and Human

    Development.

National Research Council.(1998). *Preventing reading difficulties in young children*.

    Washington, DC: National Academy Press.

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

Nettles, S.M., & Herrington, C. (2007).Revisiting the importance of the direct effects of school leadership on student achievement: The implications for school improvement policy. *Peabody Journal of Education, 82*(4), 724-736.

Nevitt, J., & Hancock, G. R. (2004). Evaluating small sample approaches for model test statistics in structural equation modeling. *Multivariate Behavioral Research, 39*, 439-478.

Newman, I. & Newman, D.O. (1994).*Conceptual statistics for beginners*.   Lanham, MD: University Press of America, Inc.

Newman, I., Newman, C., Brown, R. & McNeeley, S. (2006). *Conceptual statistics for beginners*. Lanham, MD: University Press of America, Inc.

Newman, I., Salzman, J. A., & Lenhart, L. A. (2004).Reading First Ohio Program Monitoring Tool (RFPMT). In J. A. Salzman& S. A. Brown (Eds.), *Reading First Ohio program monitoring guide and Program Monitoring Tool*. Cleveland, OH: Reading First Ohio Center.

Newmann, F. & Wehledge, G. (1995).*Successful school restructuring*. Madison, WI: Center on Organization and Restructuring of Schools.

Northwest Regional Educational Laboratory, WA. (2001). *Catalog of School Reform Models. Resources*. Retrieved (2009, August 7) from http://educationnorthwest.org

Obama, B. (2009). Address to Congress. (2009). *Presidential address to congress*. Washington, DC: United States.

Ohio Department of Education (2003).*Reading First Ohio Assurances*.

O'Neil, J. (2000). Fads and fireflies: the difficulties of sustaining change. *Educational Leadership*.6-10.

Osborne, J.A. (Ed.). (2008). *Best practices in quantitative methods*. Thousand Oaks, CA: Sage.

Pedhazur, E.J., & Schmelkin, L.P. (1991).*Measurement, design and analysis*. Hillsdale, NJ: Lawrence Earlbaum Associates.

Pogrow, S. (2001).Avoiding comprehensive school reform models. *Educational Leadership, 58*(8), 82-83.

Pogrow, S. (2000). Success for all does not produce success for students. *Phi Delta Kappan, 82*(1), 67-80.

Ravitch, D. (2000). *Left back: A century of battles over school reform*. New York: Simon and Schuster.

Robinson, V.M., & Timperley, H.S. (2007). The Leadership of the improvement of teaching and learning: Lessons from initiatives with positive outcomes. *Australian Journal of Education, 51*(3), 247-262.

Rogers, E.M. (2003). *Diffusion of innovation*. New York: Free Press.

Rosenblum, S., & Louis, K. (1979).*Stability and change: Innovation in an educational context*. Cambridge, MA: ABT Associates

Rosenholtz, S.J. (1989). *Teachers' workplace: the social organization of schools*. New York: Longman

Salzman, J.A. (2006). Leading first: classroom walkthrough (CWT) training to help principals lead their Reading First schools. *Principal Navigator, 1*(4) 10-14.

Satorra, A., & Bentler, P.M. (2001).A Scaled difference chi-square test statistic for

    moment structure analysis. *Psychometrika, 66*(4), 507-514.

Schermerhorn, J., Hunt, J., & Osborn, R. (2003).*Organizational behavior*. Hoboken, NJ:

    John Wiley and Sons.

Sebring, P.B., & Bryk, A.S. (2000).School leadership and the bottom line in Chicago. *Phi*

    Delta Kappan, 81(6), 440-443.

Senge, P.M. (1990). *The Fifth discipline: The art & practice of the learning organization*.

    New York: Doubleday.

Smith, M.W. & Dickinson, D.K. (2002).*User's guide to the Early Language and Literacy*

    Classroom Observation toolkit. Baltimore, MD: Brookes.

Smith, L., Maxwell, S., Lowther, D., Hacker, D., & Bol, L. (1997).Activities in schools

    and programs experiencing the most, and least, early implementation successes.

    School Effectiveness and School Improvement, 8(1), 63-73.

Spillane, J. (2000). Cognition and policy implementation: District policymakers and

    reform of mathematics education. *Cognition and Instruction, 18*(2), 141-179.

Spillane, J.P. & Thompson, C.L. (1997). Reconstructing conceptions of local capacity:

    The local education agency's capacity for ambitious instructional reform.

    Educational Evaluation and Policy Analysis, 19(2), 185-203.

Spillane, J., & Zeuli, J. (1999). Reform teaching: exploring patterns of practice in the

    context of national and state mathematics reform. *Cognition and Instruction,*

    21(1), 1-27.

Stringfield, S. (2002). Big change questions: Is large-scale educational reform possible?

    Journal of Educational Change, 3, 63-73.

Stringfield, S., & Datnow, A. (1998).Scaling up school restructuring designs in urban schools. *Education and Urban Society, 30*(3), 269-276.

Stringfield, S., & Ross, S.M. (1997). A Reflection at mile three of a marathon: the Memphis restructuring initiative. *School Effectiveness and School Improvement, 8*(1), 151-161.

Timperley, H.S., & Parr, J.M. (2005).Theory competition and the process of change. *Journal of Educational Change, 6*, 227-251.

Thornton, B., Shepperson, T., & Canavero, S. (2007).A Systems approach to school improvement program evaluation and organizational learning. *Education, 128*(1), 48-56.

Tyack, D., & Cuban, L. (1995).*Tinkering toward utopia: A century of public reform.* Cambridge, MA: Harvard University Press.

Tyack, D., & Tobin, W. (1994). The Grammar of schooling: why has it been so hard to reform? *American Educational Research Journal, 31*(3), 453-479.

United States Department of Education, Office of Reading First (2002).*Reading first guidance and legislation.*www2.ed.gov/programs/reading first/legislation.html

Viadero, D. (1999, June 23). Who's in, who's out? *Education Week*, 12-13.

Weatherill, K.S., & Applefield, J.M. (2005). Using School change states to analyze comprehensive school reform projects. School *effectiveness and school improvement, 16*(2), 197-215.

Whitaker, T. (2004).*What Great principals do differently.* Larchmont, NY: Eye on Education.

Whitaker, T. (2004).*What Great teachers do differently*. Larchmont, NY: Eye on

    Education.

Yuan, K. H., & Bentler, P. M. (1997). Mean and covariance structure analysis:

    Theoretical and practical improvements. *Journal of the American Statistical*

    Association, 92, 767-774.