5-2012

Mathematics achievement : the impact of America's choice in Kentucky's schools.

Sonia James Upton
University of Louisville

Follow this and additional works at: http://ir.library.louisville.edu/etd

Recommended Citation
https://doi.org/10.18297/etd/1477

This Doctoral Dissertation is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. This title appears here courtesy of the author, who has retained all other copyrights. For more information, please contact thinkir@louisville.edu.
MATHEMATICS ACHIEVEMENT: THE IMPACT OF AMERICA'S CHOICE IN KENTUCKY'S SCHOOLS

By

Sonia James Upton
B.S., Western Kentucky University, 1982
M.A., Western Kentucky University, 1987
English as a Second Language Endorsement, 1989

A Dissertation
Submitted to the Faculty of the
College of Education and Human Development
of the University of Louisville
In Cooperation with Western Kentucky University
In Partial Fulfillment of the Requirement
For the Degree of

Doctor of Philosophy

Department of Leadership, Foundations and Human Resource Education
University of Louisville
In cooperation with
Western Kentucky University

May 2012
MATHEMATICS ACHIEVEMENT: THE IMPACT OF AMERICA’S CHOICE IN KENTUCKY’S SCHOOLS

By

Sonia James Upton
B.S., Western Kentucky University, 1982
M.A., Western Kentucky University, 1987
English as a Second Language Endorsement, 1989

A Dissertation Approved on

April 16, 2012

by the following Dissertation Committee:

Dr. Stephen K. Miller
Dissertation Chair

Dr. Kyong Hee Chon
Methodologist

Dr. Jennifer M. Bay-Williams

Dr. Elizabeth B. Patton

Dr. Bud Schlinker
DEDICATION

This dissertation is dedicated to

God, my constant guide, companion, and strength,

through Him all things are possible,

and

Jason Alan James, my son and friend

who encouraged and supported me during my coursework,

and

Norman Devoye Upton, my husband, friend, and companion

who gave me invaluable support and encouragement during the writing process.
ACKNOWLEDGEMENTS

The completion of a dissertation is not an easy task and is in no way accomplished without the patience, support, understanding, and encouragement of many people. Without these elements in place, it is hard to reach the end successfully. I am forever indebted to those who were there for me along this journey.

First of all, I would like to thank God who blessed me with determination, drive and persistence to complete any task and achieve my goals no matter what odds or obstacles may appear in my path.

Secondly, I would like to thank Dr. Stephen Miller, my dissertation chair, for his dedication, guidance, knowledge and leadership. His patience and support was the anchor that kept me on course and moving towards completion. He patiently guided me along the path of discovering the writer hidden within. I would also like to thank my methodologist, Dr. Kyong Hee Chon for her statistical support and knowledge that added strength to my study. I also owe a debt of gratitude to my other committee members Drs. Bud Schlinker, Elizabeth B. Patton, and Jennifer Bay-Williams for their valuable time and input throughout the writing process.

Thirdly, I owe much gratitude to family and friends who supported and encouraged me throughout the various stages of the doctoral process. Thanks for your patience and understanding as I retreated from activities and spent time working.

Last but not least, I would like to thank all of my English language learners who entered my classroom. They inspired me to always soar higher and begin this journey.
This study examined student achievement scores in Kentucky elementary schools to determine the relationship between implementing the America’s Choice comprehensive school reform model and student achievement in mathematics. Six research questions guide this study: For the seven America’s Choice schools for third grade mathematics: (1) To what degree do the America’s choice schools differ from the statewide mean for math achievement at the beginning of the program, (2) To what extent is the year of implementation related to math achievement, (3) To what extent are differences between schools related to math achievement, (4) To what extent is the beginning year of the program related to math achievement, (5) What is the effect of controlling for demographic factors on Research Questions 2-4, and (6) For the three years of the America’s choice mathematics program (Year 1, Year 2, and Continuation 1), how does the progress of the America’s Choice schools compare to that statewide for the same period?

For this study, the researcher analyzed the Comprehensive Test of Basic Skills (CTBS/5) Normal Curve Equivalency (NCE) mathematics scores for third grade/exiting primary students in seven Kentucky schools implementing America’s Choice in
mathematics. Achievement data from the Kentucky Department of Education for the years 2001-2005 were collected across four cohorts, including year prior to implementation, two years of implementation, the following continuation year, and related demographic factors. Research Questions 2-5 utilized two multiple regression models to study the relationships between math achievement, change model predictors (year of implementation, schools, beginning year), and demographics. One sample \textit{t} tests were employed for RQs 1 and 6. All analyses were conducted using the Statistical Package for the Social Sciences (SPSS) computer program.

Results of the study on the seven Kentucky elementary schools were inconclusive. While the statistical analysis indicated gains in mathematics achievement for some schools, others demonstrated declines during implementation years and continuation year. For demographic variables, size of school had the strongest correlation \((r = .140)\) with NCE math scores.

This research provides insights for schools considering adopting America's Choice as a reform model impacting mathematics achievement. Since this study was specific to Kentucky, future research is needed to expand to other states.
# TABLE OF CONTENTS

DEDICATION ................................................................................................................... iii  
ACKNOWLEDGEMENTS ............................................................................................... iv  
ABSTRACT ........................................................................................................................ v  
TABLE OF CONTENTS .................................................................................................. vii  
LIST OF FIGURES ............................................................................................................ x  
LIST OF TABLES .............................................................................................................. xi  

**I. STATEMENT OF THE PROBLEM.** ........................................................................ 1  
  Introduction ..................................................................................................................... 1  
    Kentucky Initiatives ........................................................................................................ 2  
    School Reform ................................................................................................................ 6  
    The Problem Defined ..................................................................................................... 12  
    Purpose of Study ............................................................................................................ 14  
    General Lineage of Reform Framework ....................................................................... 15  
    Research Questions ....................................................................................................... 16  
    Significance of Study ................................................................................................... 17  
    Assumptions and Limitations ....................................................................................... 18  
    Definition of Terms ...................................................................................................... 20  
    Summary ....................................................................................................................... 21  

**II. REVIEW OF LITERATURE.** .............................................................................. 23  
  Introduction ..................................................................................................................... 23  
  School Reform ................................................................................................................. 25  
  Comprehensive School Reform ...................................................................................... 33  
  America’s Choice School Reform .................................................................................. 54  
  Mathematics Achievement ............................................................................................. 65
Summary............................................................................................................... 84

III. METHODOLOGY.................................................................................................. 87
   Introduction............................................................................................................. 87
   Study Design........................................................................................................... 88
   Population and Sample......................................................................................... 91
   Data Analysis.......................................................................................................... 96
   Description of the Variables.............................................................................. 105
   Validity and Reliability....................................................................................... 108
   Ethical Considerations....................................................................................... 109
   Summary............................................................................................................. 109

IV. RESULTS ............................................................................................................. 111
   Introduction......................................................................................................... 111
   Descriptive Statistics.......................................................................................... 112
   Research Questions............................................................................................. 119
   Summary............................................................................................................. 137

V. DISCUSSION AND CONCLUSION...................................................................... 141
   Introduction......................................................................................................... 141
   The Study in Brief............................................................................................... 141
   Discussion.......................................................................................................... 143
   Recommendations............................................................................................... 164
   Conclusions......................................................................................................... 171

REFERENCES ........................................................................................................... 176
APPENDIX A........................................................................................................... 189
APPENDIX B........................................................................................................... 191
APPENDIX C........................................................................................................... 192
APPENDIX D........................................................................................................... 195
APPENDIX E........................................................................................................... 198
CURRICULUM VITAE............................................................................................. 200
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Comparing Kentucky’s SISI to America’s Choice School Design.</td>
<td>11</td>
</tr>
<tr>
<td>2. General lineage of levels of School Reform</td>
<td>16</td>
</tr>
<tr>
<td>3. Sequence of America’s Choice Mathematics Implementation over time.</td>
<td>16</td>
</tr>
<tr>
<td>4. Comparing weighted %Change Scores for America’s Choice versus the State</td>
<td>137</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Years of Implementation of the America’s Choice Model in Kentucky Schools</td>
<td>89</td>
</tr>
<tr>
<td>2. Descriptive Characteristics for America’s Choice Schools in Kentucky for 2000</td>
<td>90</td>
</tr>
<tr>
<td>3. America’s Choice Components Implemented by Each School</td>
<td>91</td>
</tr>
<tr>
<td>4. Years of Implementation of America’s Choice Mathematics in Kentucky Schools</td>
<td>92</td>
</tr>
<tr>
<td>5. Number of Third Graders in Each School during Years of the Study</td>
<td>94</td>
</tr>
<tr>
<td>6. Descriptive Statistics for Demographic Variables Size and Percent Free/Reduced Lunch for Third Grade across the Seven Schools</td>
<td>114</td>
</tr>
<tr>
<td>7. NCE Grade 3 Mathematics Scores for Seven America’s Choice Schools across the Four Implementation Years</td>
<td>117</td>
</tr>
<tr>
<td>8. Variables, Variable Label Codes, and Type of Data Utilized</td>
<td>118</td>
</tr>
<tr>
<td>9. One Sample t test Comparing Means for the Seven Schools to the State Mean for Prior Year</td>
<td>120</td>
</tr>
<tr>
<td>10. Hierarchical Regression of Mathematics NCE Scores on Three Blocks of America’s Choice Implementation Model</td>
<td>123</td>
</tr>
<tr>
<td>11. Hierarchical Regression of Mathematics NCE Scores on Demographic Variables and Three Blocks of America’s Choice Implementation Model</td>
<td>128</td>
</tr>
<tr>
<td>12. One Sample t Tests Comparing Mean NCE Third Grade Mathematics Scores for America’s Choice Versus the State for Year 1, Year 2, and Continuation Year</td>
<td>132</td>
</tr>
<tr>
<td>13. Mean NCE Math Scores Utilized for Cohort Percentage Change Calculations for America’s Choice and Statewide</td>
<td>134</td>
</tr>
</tbody>
</table>
14. Percentage Change Scores across Staggered Start Dates for Year 1, Year 2, and Continuation Year for America’s choice Schools and Statewide ............................. 135

15. Weighted Percent Change Scores for Seven America’s Choice Schools and for Statewide for Year 1, Year 2, and Continuation Year. ........................................ 136

E-1. Pearson $r$ Correlation Matrix for NCE Mathematics Scores, and Demographic Variables. ........................................................... 199
CHAPTER I

STATEMENT OF THE PROBLEM

Introduction

The launch of Sputnik can be linked as one cause of the United States federal government's large-scale nationwide curriculum reform initiatives towards the end of the 1950s and early 1960s. Sputnik's orbit raised concerns in the United States that the Soviet school system was superior to America's school system. The government's goal was to develop initiatives that would reform the educational system in the United States and yield the desired results of stronger mathematics and science curriculum in addition to strong literacy curriculum.

Early educational reform movements during the 1960s and 1970s aimed at decentralization and involved political or administrative agendas (David, 1989). As part of President Lyndon Johnson's war on poverty, Congress enacted the Elementary and Secondary Education Act (ESEA) on April 11, 1965. This was the most expansive federal education bill authorizing funds for professional development, instructional materials, resources to support educational programs, and promotion of parental involvement. The call to action inherent in the Coleman Report (Coleman et al., 1966) on poverty, race, and inequality; the threat from Sputnik; and prior presidents' commitment to improving America's educational system—all helped inspire the passage of this law. The ESEA Act was originally authorized through 1970 and has been reauthorized by the government in various forms since its original enactment. The current reauthorization of
ESEA is the No Child Left Behind Act of 2001.

During the 1980s, citizens became concerned that the public educational system in America was not adequately preparing youth to be successful members of society upon graduation. This growing concern led to the development of educational initiatives with a stronger focus on school accountability for teachers and instruction (Fullan, 1991; Lee, 2006; Murphy, 1990). Educational reform movements during 1980–1987 focused on school accountability. Policy makers became concerned that U.S. students were falling behind those in other countries, yet compared to their foreign counterparts, U.S. schools remained subject to little accountability. Bureaucratic top down initiatives focused on instruction and achievement through mandated curricular changes. Despite this movement, schools still had many low performing students, particularly for at-risk populations in the areas of reading and mathematics.

There followed a mini-cycle of teacher empowerment (Murphy, 1990) but this too failed to bring about the desired changes. The Agenda for Action (National Council of Teachers of Mathematics [NCTM], 1980), which helped shape some of the changes in public education, focused on improving students’ reasoning, thinking, and problem solving skills in reading and mathematics (Bailey, 1980). During 1988-1995, reform movements focused on school choice, professional sanctions, and decentralization (David, 1989; Heck, Brandon, & Wang, 2001; and Marks & Louis, 1999). Ultimately all of these brief cycles were replaced by the standards-based curricular reforms of the 1990s which continue to impact today’s policy changes (Fullan, 1991; Goertz & Duffy, 2003; Murphy, 1990).

Kentucky Initiatives

Kentucky was one of many states that had their public school financing systems
invalidated as a result of finance equity suits. Kentucky’s court ruling (*Rose v. Council*, 1989) made the distinction that the *entire* system of common schools was unconstitutional and ordered the legislature to recreate or reestablish a new public education system. In response to this court decision, the Kentucky General Assembly passed into law the Kentucky Education Reform Act of 1990 (KERA, House Bill 940), widely recognized as the most comprehensive and ground breaking among the 50 states (Pankrantz & Petrosko, 2000; Steffy, 1993).

The purpose of KERA was to revamp completely Kentucky’s educational system in the areas of finance, governance, and curriculum in an effort to provide equal educational opportunities for all of Kentucky’s children. Through the implementation of KERA, learning standards were established for all students. The new standards were related to the National Assessment of Educational Progress (NAEP) and were significantly higher than the state’s previous educational standards (Jencks & Phillips, 1998; Petrosko, 2000; Rothstein, 2004). The Kentucky standards have guided reform efforts since KERA was implemented in 1990. Improvements in curriculum and instruction as well as changes in the accountability system all are based upon the following six goals and academic expectations, which have remained unchanged since their original formulation. According to the current citation (Kentucky Department of Education, 2007, ¶ 4), students shall develop:

1. basic communication and mathematics skills for purposes and situations they will encounter throughout their lives.

2. their abilities to apply core concepts and principles from mathematics, sciences, arts, humanities, social studies, practical living studies, and vocational studies to what they will encounter throughout their lives.

3. their abilities to become self-sufficient individuals
4. their abilities to become responsible members of a family, work group, and community, including demonstrating effectiveness in community service.

5. their abilities to think and solve problems in school situations and in a variety of situations they will encounter in life.

6. their abilities to connect and integrate experiences and new knowledge from all subject matter fields with what they have previously learned and build on past learning experiences to acquire new information through various media services.

Kentucky Reform Efforts

The Kentucky Education Reform Act (KERA) constituted systematic reform; the focus, broader than most states, included several essential components: professional development, preschool programs, primary programs, family resource and youth service centers, extended school services, technology, curriculum development, assessment, school-based decision making, and regional service centers (Brown & Warschauer, 2005; DeYoung, 1994; Ennis, 2002; McKinney, 2007; Pankratz & Petrosko, 2000; Steffy, 1993). The guiding principle of the KERA reform was that all children could learn at high levels, given sufficient time, effort and opportunity (Ennis, 2007; Saravia, 2008; Steffy, 1993). A number of innovations were introduced to Kentucky's educational system as part of the implementation of KERA to ensure that public education would (a) provide resources equitably across all school districts, (b) provide resources to districts with large numbers of disadvantaged children, (c) eliminate political favoritism, (d) set high standards for everyone involved in public education, (e) provide a technology support network, (f) empower local schools to make their own decisions to support education efforts, (g) hold schools accountable by setting standards, (h) reward successful schools, and (i) assist unsuccessful schools (Kentucky Department of Education, 2000a).

As part of the school improvement process, Kentucky extended standards-based
accountability from curriculum content to whole school reform. The *Standards and Indicators for School Improvement: Kentucky's Model for Whole School Improvement* (SISI), adopted by the Kentucky Board of Education, represent a framework for measuring a school’s readiness to increase student achievement (Kentucky Department of Education, 2004a). The SISI document consists of nine standards, each with its own set of research-based indicators, organized into three sections: Academic Performance, Learning Environment, and Efficiency. A complete list of the nine *Standards and Indicators for School Improvement* appears in Appendix D.

As part of the No Child Left Behind legislation, regulations were established for schools that did not make adequate yearly progress (AYP) based on the Kentucky Core Content Tests (KCCT) items in reading, mathematics, and other non-cognitive criteria. A school’s consequence for not making AYP included a Scholastic Audit, receiving additional funds, revising school improvement plans, assistance from a Highly Skilled Educator, an evaluation of school personnel, and transfer of students to other more successful schools (Kentucky Department of Education, 2003).

The Scholastic Audit process uses the *Standards and Indicators for School Improvement* (Kentucky Department of Education, 2004a) to assess a school’s capacity for improvement based on the nine indicators. This audit provides diagnostic information for schools as the staff works toward their yearly goals to attain proficiency by 2014. Schools are encouraged to undergo a scholastic review using the SISI document as a means of self-assessing to determine if they are on track towards meeting their goals. Schools with assessment scores in the lowest one-third were classified as Tier 3 Assistance. Schools with scores in the middle were classified as Tier 2 Assistance and schools with scores in the top third were classified as Tier 1 Assistance.
According to NCLB Section 1116 (c) (9), Kentucky schools that are in Tier 3 Assistance for failure to meet their achievement goals must undergo the Scholastic Audit process. Mintrop (2003) praised the Scholastic Audit as a powerful reform mechanism when used as a model of support for struggling schools. In response to the Scholastic Audit findings, some schools adopt and implement various comprehensive school reform models utilizing Title I and supplemental funds. These funds are available to assist schools to make improvements with support from the state.

School Reform

As just described, the innovations implemented under KERA during the 1990s were both extensive and in several instances, ground-breaking. Many schools in Kentucky began to show progress. Within the larger context of reform nationally, however, the No Child Left Behind Act (NCLB), the latest reauthorization of the 1965 Elementary and Secondary Education Act, was passed by the U.S. Congress in 2001 and signed into law by President George W. Bush in 2002 (NCLB, 2002). NCLB and the Comprehensive School Reform Demonstration Program (CSRD) of 1998 were developed to assist schools that were still in need of assistance to improve their educational programs.

Comprehensive school reform designs were intended to reorganize and direct numerous elements of the educational experience in an effort to improve student achievement. The implementation of comprehensive school reform focused on nine critical factors: Characteristics of change (Need, Clarity, Complexity, and Quality/Practicality), Local Characteristics (District, Community, Principal, and Teacher), and External Factors (Government and other agencies). Several researchers have concluded that in order for students to experience sustainable growth in
Mathematics, comprehensive standards, research-based curriculum, and a change in daily instruction must be implemented (Brown & Woodward, 2006; Desimone, 2002; Erlichson, 2005; Schoenfield, 2002; Supovitz & Poglinco, 2001).

*Mathematics Achievement in School Reform*

The National Council of Teachers of Mathematics (NCTM) has been a leader in the development of guiding principles that helped shape the mathematics reform movement. Such resources as *An Agenda for Action: Recommendations for School Mathematics* (NCTM, 1980), *Curriculum and Evaluation Standards for School Mathematics* (NCTM, 1989), *Professional Standards for Teaching Mathematics* (NCTM, 1991), *Assessment Standards for School Mathematics* (NCTM, 1995), and *Principles and Standards for School Mathematics* (NCTM, 2000) called for math teachers to move away from teaching by telling (the traditional approach) and placed greater emphasis on real-world problem solving, building connections, and developing understanding of math concepts and skills (the constructivist teaching paradigm). The *Principles and Standards for School Mathematics* (NCTM, 2000) provides the framework for the development of strong mathematics programs that support the development of critical thinking skills.

The National Assessment of Educational Progress (NAEP) evaluates academic achievement by assessing samples of students representing the nation. Mathematics assessments are administered to students in grades 4, 8, and 12. The NAEP assessments are scored and reported in the Nation's Report Card according to three achievement levels: Basic (partial mastery), Proficient (solid academic performance), and Advanced (superior performance). The percentages of fourth-graders performing at or above Basic (82 percent) and at or above Proficient (39 percent) in 2009 were unchanged from those in 2007. The percentages of eighth-graders performing at or above Basic (73 percent) and
at or above Proficient (34 percent) in 2009 were higher than those of previous assessment periods (NAEP, 1990). Trends for the preceding cycle showed steady increases from year to year. Even though the mathematics trends increased, there was still concern that U.S. students were lagging behind other nations (Fullan, 1991; Kenny & Silver, 1998; Lee, 2006).

As mathematics becomes more important around the world, educators seek ways to increase student understanding of mathematical skills and concepts while also raising student achievement scores. The instructional strategy of repeated practice of computational algorithms must shift to a focus on emphasizing meaningful experiences in mathematics classrooms (NCTM, 1989). This change in focus of instruction called for a drastic restructuring of traditional mathematics curricula and assessment practices. Assessment procedures must no longer deny students the opportunity to learn important mathematics; instead assessment should be a means of fostering growth towards high expectations (Tate, 1997; Van Haneghan, Pruet, & Bamberger, 2004). The mathematics workshop component of America’s Choice comprehensive reform focuses on providing differentiated instruction to help students develop basic skills, problem solving, and conceptual understanding.

**Comprehensive Reform in Kentucky**

In 1994, Congress changed the focus of Title I programs under the Improving America’s Schools Act (IASA). This law, which reauthorized the Elementary and Secondary Education Act (ESEA), included Title I provisions calling for schools that receive Title I funds to set high standards for all students, to assess all students relative to these standards, to report results to the public, and to make instructional and structural changes to ensure that all students have the opportunity to meet these standards (Padilla et
This focus came as program evaluations suggested that targeted, “pull-out” education programs for students, the previous strategy for Title I funds, showed no clear positive effect on student achievement in high poverty schools. Now the use of Title I, Part A grant funds has been expanded to provide for a range of assistance including data analysis, school improvement planning, and curriculum alignment. Over the years, federal and state policy makers have mandated that schools that are low performing based on their state achievement testing system implement a comprehensive school reform design to address those needs. This initiative led to the development of several comprehensive school-wide reform models implemented utilizing Title I funds.

Kentucky schools selected from a variety of comprehensive reform models based on their identified need. The models Kentucky schools implemented included Accelerated Schools Project, America’s Choice, ATLAS Communities, Early Intervention in Reading, First Steps, Modern Red Schoolhouse, School Development Program, and Success for All (St. John, Loescher, Jacob, Cecki, Kupersmith, & Musoba, 2000).

*America’s Choice*

Implementation of America’s Choice reform follows a three year process. During the first year, schools focus on reading instruction through Reader’s Workshop and writing instruction through Writer’s Workshop. The second year of implementation, schools continue with Reader’s and Writer’s Workshops and add Mathematics Workshop. Science is added during the third year of implementation (Allen, Knight, & Matthews, 2003). During this time, schools receive layered assistance and support for implementation from America’s Choice staff. As part of the assistance, America’s Choice staff provided training for literacy and mathematics coaches in each school, with continued support and training.
from America’s Choice staff lasting for three years.

Improved mathematics instruction is a goal set by many schools and districts in the United States. The same is true for Kentucky schools who are finding it difficult to reach the goal of proficiency by the 2013–2014 school year. Many schools in Kentucky that have high poverty status have adopted comprehensive models in order to address this issue, under funding provided by Title I.

As noted, the America’s Choice comprehensive school reform was among those implemented in Kentucky elementary schools to address long term achievement goals, specifically the levels of proficiency for all students by the 2013-2014 school year. America’s Choice is a comprehensive reform model designed by the National Center on Education and the Economy (NCEE) for grades K through 12. This design focuses on raising academic achievement by providing a rigorous standards-based curriculum and safety net for all students. The elements of America’s Choice are grouped into five components: high performance leadership and management; capacity building through aligned instructional systems; professional learning communities; improved student achievement for at-risk learners through standards and assessments; and parent and community involvement (Allen, Knight, & Matthews, 2003). Schools or districts selecting to implement this reform model usually have a history of low student achievement.

One hallmark of comprehensive school reform has been the emphasis on empirical evaluation of the change efforts. Kentucky schools in Tier 3 Assistance were required to adopt a research-based reform model; America’s Choice was one such comprehensive program (National Center on Education and the Economy [NCEE], 2002).

After reviewing results from their Kentucky Department of Education (KDE) sponsored Scholastic Audit, each school reviewed several reform options before selecting a
specific model. From 2000 through 2004, nine elementary schools across Kentucky opted for America’s Choice to guide schoolwide improvement and move their school out of assistance. *America’s Choice School Design, a Good Choice for Kentucky* (Allen et al., 2003) explains how Kentucky’s *Standards and Indicators for School Improvement* (SISI) aligns with America’s Choice comprehensive reform components (See Figure 1). Each of the three sections of the *Standards and Indicators for School Improvement* (Academic Performance, Learning Environment, and Efficiency) defines elements of school-wide improvement. These same elements are addressed in the America’s Choice reform design.

<table>
<thead>
<tr>
<th>Standards and Indicators for School Improvement</th>
<th>America’s Choice Comprehensive Reform Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Performance Standards 1, 2, 3</td>
<td>Focus on standards, assessment, and aligned instructional systems. Performance standards provide examples of student work. Includes alignment between state standards, state assessments, and instructional practices. Specific examples describe how the design helps a school use standards, plan instruction and evaluate student work against a standard.</td>
</tr>
<tr>
<td>Focus on curriculum, instruction, and classroom evaluation and assessment.</td>
<td></td>
</tr>
<tr>
<td>Learning Environment Standards 4, 5, 6</td>
<td>Helps schools address learning environment issues. Teachers and students learn the rituals and routines for the Readers, Writers, and Math Workshops and classroom practices. Staff receives intensive training through a variety of sources to learn to implement best practices. Includes support for working with the Parent Community Outreach Coordinator to include community as a resource.</td>
</tr>
<tr>
<td>Focus on school culture, student family, and community support and professional growth and development</td>
<td></td>
</tr>
<tr>
<td>Efficiency Standards 7, 8, 9</td>
<td>Major components include management, high performance leadership, and organization. Leadership receives training through Principal Academies and scheduled network meetings. Planning for results provides specific examples for using the design with Kentucky’s Comprehensive School Improvement Planning process.</td>
</tr>
<tr>
<td>Focus on leadership, organizational structure and resources, and comprehensive and effective planning.</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 1. Comparing Kentucky’s SISI to America’s Choice comprehensive reform design.*
America's Choice works with districts and schools to align classroom instruction with state standards and assessments. The ongoing analysis of student assessment data and student work allows teachers focus instruction on identified needs and move students toward attainment of standards. During the mathematics workshop, the teacher introduces a math concept to the students. This is followed by a forty minute period that allows the students to explore and develop understanding of the new concept. The workshop concludes with an opportunity for the students to share and explain their understanding. This format provides opportunities for students to develop critical thinking skills and develop deeper understanding which aligns with the focus and direction of the NCTM standards.

The Problem Defined

In the standards-based accountability era (cf. Fullan, 1991; Murphy, 1990), several trends can be observed. First, accountability models dominate current approaches to improving schools with respect to the overwhelming achievement gaps between at-risk and middle class students by having the same high standards for the disadvantaged as for their affluent peers (cf. Goertz & Duffy, 2003). In many respects these "blaming the victim" arguments have shifted from the students to the teachers who are now held accountable for achievement for all students equally regardless of their class or racial background (McDermott, 2007; Ryan, 1976).

Second, much of the accountability work is disciplined-based, with the greatest focus on mathematics and English (some combination of reading and writing). The work in these fields (as well as science and social sciences) revolves around standards-based curriculum and instruction. In the area of mathematics, these debates often involve a major shift in the instructional practices, from traditional methods of lecture, explanation,
and homework practice on information presented by the teacher to constructivist models in which students become more active in their learning as they internalize (construct) their own mathematical insights and understanding (Gregg, 1995; Hiebert, 1992; Hiebert et al., 1998; Larochelle, Bednarz & Garrison, 1998; Noddings, 1993; Simon, 1995).

Complicating this movement is the fact that traditional instruction dies hard (cf. Cuban, 1990) and teachers prove resistant to the student-centered approach.

Third, comprehensive reform researchers have ascertained that models must be comprehensive and systematic to be successful in bringing about effective school-wide change (Ball & Cohen, 1996; Borman, Hewes, Overman, & Brown, 2003; Dimmock, 2002; Fullan, 1991). Several comprehensive models have been developed (e.g., Accelerated Schools Project, America's Choice, ATLAS Communities, Early Intervention in Reading, First Steps, Modern Red Schoolhouse, School Development Program, and Success for All) and despite specific differences, they all have in common the systematic, school-based, comprehensive foundation. Title I funding became a mechanism for getting low performing schools the assistance that was needed in order to undertake such an ambitious project.

Within the context of these broader trends, Kentucky supported Tier 3 schools (not reaching long term achievement goals) in selecting one of several comprehensive reform models. For nine Kentucky schools, America's Choice comprehensive reform was intended to produce changes in the content areas of reading, writing, mathematics and science. VanMeter (2005) examined the effects of implementing America's Choice for third grade reading achievement in nine elementary schools in Kentucky. To date no one has conducted a comparative assessment of these nine schools in the area of third grade mathematics achievement. Therefore the question remains about the lasting or sustaining
effects of mathematical success after the initial two years of implementation of this design. The research problem for this study, therefore, is whether schools that have implemented the America's Choice comprehensive reform model for school improvement continue their level of mathematics achievement gains after the initial two years of grant-funded implementation have ended.

Purpose of Study

This study brings together the issues noted in the Problem Defined. First, since KERA was implemented in 1990, Kentucky's overall performance has been driven by a value-added accountability model (cf. Miller, 1992; Steffy, 1993). The Kentucky accountability model expressly forbids taking account of demographic differences under the guise that all students can achieve at the same high levels, with all schools expected to reach the proficiency level by 2014, regardless of differences in their initial baseline scores.

Second, mathematics educators continue in their efforts to change instruction toward more active, constructivist practices (Hiebert, 1992; Kenny & Silver, 1998; Reys et al., 1997; Van Haneghan et al., 2004). The persistence of traditional math instruction is often exacerbated in the context of school accountability because of pressures to teach to the test. The broad context of statewide performance testing raises concerns that are not necessarily addressed in many of the micro studies on mathematical instructional strategies and curriculum reform (Elliott, 1996; Henry & Clements, 1999).

Third, there is a continuing need for additional research on the effects of comprehensive school reform efforts. This is complicated because of differences from one reform package to the next. In this regard, distinguishing the differential effects of implementation of comprehensive school reform generally (e.g., KERA statewide reform
and improvements) from the effects of the various reform models can be difficult. Then too, the specifics of comprehensive reform can vary from state to state. Because of the unique conditions in Kentucky, acknowledged as among the most extensive and wide-ranging reform packages (Pankrantz & Petrosko, 2000), research in that state may provide contextual insights about the dynamics of implementation. Finally, there is the need to extend VanMeter’s (2005) work on America’s Choice in Kentucky (third grade reading) to mathematics. Either directly or indirectly, this research addresses all of these issues.

Specifically, this research builds on the work of VanMeter (2005) who examined the effects of America’s Choice comprehensive reform for reading in nine elementary schools in Kentucky. Utilizing quantitative techniques, the current study examined America’s Choice comprehensive school reform design on mathematics achievement in those same Kentucky elementary schools. Specifically, the research examined three dimensions of the implementation of America’s Choice school design for mathematics instruction for third grade exiting primary students: differences across schools; the effects of implementation over time, including the extent of growth in academic achievement in the content area of mathematics after the layers of support for the implementation of the program have been discontinued; and the staggered start date across schools. Thus the central research question for this study is; what are the effects of implementing America’s Choice comprehensive school reform design in mathematics in seven elementary schools in Kentucky?

General Lineage of Reform Framework

The general reform framework for this quantitative study is based on policy linking broader school reform to successively more forward nested strategies that are
played out in a specified discipline such as mathematics at the classroom level. Figure 2 displays the lineage of these levels of reform for this study.

![Diagram of School Reform Movements]

Figure 2. General lineage of levels of School Reform.

The sequence of implementation for America’s Choice for elementary schools in Kentucky is depicted in Figure 3. The analysis is designed to determine the impact of America’s Choice comprehensive reform on mathematics student achievement. In contrast to the reading emphasis (see VanMeter, 2005), the schools implementing this model did not add mathematics until their second year of implementation. Thus, the analysis has only two years of supported reform compared to three years for reading.

<table>
<thead>
<tr>
<th>Achievement in Year Prior to Implementation</th>
<th>Achievement after Year 1 Implementation of Math Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence of AC Math Implementation</td>
<td>Achievement after Year 2 Implementation of Math Component</td>
</tr>
<tr>
<td></td>
<td>Achievement for Continuation Year after Active Implementation</td>
</tr>
</tbody>
</table>

Figure 3. Sequence of America’s Choice mathematics implementation over time.

Research Questions
The researcher collected and analyzed Comprehensive Test of Basic Skills (CTBS/5) Normal Curve Equivalent (NCE) mathematics assessment data for third grade exiting primary students in the seven Kentucky elementary schools implementing the mathematics component of the America’s Choice comprehensive design. Nine schools adopted America’s Choice but only eight completed the mathematics component; seven continued the mathematics component for two years. The study addressed the following research questions:

For the seven America’s Choice schools for third grade mathematics:

1. To what degree do the America’s Choice schools differ from the statewide mean for math achievement at the beginning of the program (Prior Year)?
2. To what extent is the year of implementation related to math achievement?
3. To what extent are differences between schools related to math achievement?
4. To what extent is the beginning year of the program related to math achievement?
5. What is the effect of controlling for demographic factors for Research Questions 2-4?
6. For the three years of the America’s Choice mathematics program (Year 1, Year 2, Continuation 1), how does the progress of the America’s Choice schools compare to that statewide for the same period?

Significance of Study

The knowledge gained from this quantitative study will be beneficial to district and local Kentucky public school personnel as they consider adopting and implementing Comprehensive School Reform (CSR) models like America’s Choice. Given the goals of increasing student academic achievement and making adequate yearly progress, a part of NCLB requirements (KDE, 2004b), these results added to the existing knowledge base on school reform. Beyond these general statements, this study is significant due to the following factors.
First, this study examines mathematics achievement as part of one specific comprehensive reform model, specifically America’s Choice comprehensive school reform, and its impact on third grade exiting primary mathematics achievement scores.

Second, the current study constitutes data collected over time whereas education research generally is cross sectional. However, Kentucky’s accountability testing is a cohort model so that successive years of third graders are compared. Thus, this is not a true longitudinal design because the same students are not followed.

Third, the current study represents a conceptual replication of VanMeter’s (2005) work on reading for the same schools. Such replications are too seldom conducted in educational research.

Fourth, the current study goes beyond the VanMeter (2005) study by adding demographic controls and by investigating the impact of the America’s Choice implementation model, specifically by differences across schools, year to year improvements, and the effect of the staggered starting date.

Fifth, because KERA has been recognized for its extensive and ground-breaking set of reforms (Mintrop, 2003; Pankrantz & Petrosky, 2000; Steffy, 1993), the context for implementing various comprehensive packages is likely unique. States that have implemented comprehensive reform models may not be operating under comparable conditions, including the value-added cohort model (cf. Miller, 1992).

Assumptions and Limitations

There are two assumptions relating to this study. First, the researcher presumes that the Comprehensive Test of Basic Skills (CTBS/5) third grade exiting primary mathematics scores from the Kentucky Department of Education were valid and reliable. Second, the study presumes that all third grade mathematics scores from all of the
schools were useable for the data analysis. Treatment of records with missing data is addressed in Chapter III. Additional limitations apply; these are discussed in turn.

There are several limitations for this study. First, only nine Kentucky elementary schools implemented the America's Choice comprehensive reform design and were included in this study. One of the nine schools did not implement the America's Choice mathematics component, and one did not continue implementation for year two. Thus, this sample of seven is too small to be confident that results are stable for most schools implementing this model of reform.

Second, since only elementary data were analyzed, generalizing to middle school and high school is clearly not warranted.

Third, because KERA has been recognized for its extensive and ground-breaking set of reforms (Pankrantz & Petrosko, 2000; Steffy, 1993), the context for implementing various comprehensive reform packages is likely unique. Specifically, this suggests that the data from Kentucky regarding America's Choice comprehensive school reform model may not extend to other states.

Fourth, Kentucky's accountability model at the time of this study constituted a cohort model, with grade levels for one year compared to the same grade (thus different students) for subsequent years. Thus all results are subject to cohort variation, which can be especially problematic at the elementary level where grade level Ns are often small.

Fifth, since only mathematics data were analyzed, the findings and results do not extend to other subject areas. However, these results can be compared directly to the degree of change for reading since VanMeter (2005) analyzed the same data set for reading outcomes.

Sixth, the seven schools implementing America's Choice math did not all begin in
the same year. Under KERA, schools across the Commonwealth have collectively improved achievement levels significantly (Petrosko, 2000; Poggio, 2000) and several studies of Kentucky's Scholastic Audit data base confirm that the achievement has been rising across the elementary schools in the state from year to year (Ennis, 2007; McKinney, 2007; Saravia, 2008). Thus schools were likely at a higher level when they began implementation than they would have been if they had begun the project one or two years earlier. This staggered start is examined as part of the implementation model but individual schools may vary.

Seventh, the demographic data reported to the state of Kentucky on the Comprehensive Test of Basic Skills (CTBS/5) Normal Curve Equivalent (NCE) mathematics assessment may include inaccuracies or missing information due to student self-reporting versus school-reporting of all demographic information (including participation in the Free/reduced lunch program) on the individual student test booklets. This may have impacted the demographic data analysis included as part of the current study.

Eighth, measures of level of implementation (student survey data) were available but beyond the scope of this research. Thus, insight will be limited regarding possible explanations for why the results were found.

Definition of Terms

The following definitions are provided by the researcher to assure understanding of the terms commonly used throughout the study.

*No Child Left Behind* (NCLB) Act of 2001 (Public law 107-110) established accountability standards for schools and the state. It also established state testing requirements designed to improve education. States must categorize adequate yearly
progress (AYP) objectives and disaggregate test results for all students and subgroups based on socioeconomic status, race and ethnicity, English language proficiency, and disability (KDE, 2004b).

*Title I* is a multi-faceted Federal program providing additional funding to schools based on their ratio of students ranked at or below the poverty level (KDE, 2004c).

*School-wide Title I School* refers to schools that receive Title I Federal funds and have at least 50% of their students receiving free or reduced price lunches (U.S. Department of Education, 1996).

*Education reform* is a plan or movement which attempts to bring about a systemic change in educational practices. Educational theories, curriculum reform, and operational structure are often areas targeted for change (KDE, 2000b).

**Summary**

From the 1950s through today, the United States government has been concerned with America’s school system and the ability to provide adequate high quality educational experiences for all students to prepare them to become successful members of society upon graduation. Initiatives such as the original Elementary and Secondary Education Act (ESEA) from 1965 and more recently No Child Left Behind Act of 2001 provided directives and guidance for school districts to implement changes that would lead to the development of stronger literacy, science, and mathematics curriculum.

In 1990, Kentucky completely redesigned their educational system through the development of the Kentucky Education Reform Act (KERA) in an effort to provide high quality educational experiences to all children in Kentucky schools. Included as part of KERA, the Kentucky Board of Education adopted the *Standards and Indicators for School Improvement: Kentucky’s Model for Whole School Improvement* (SISI) (KDE,
2004a). The SISI document and the Scholastic Audit process (an in-depth analysis of each of the 88 indicators across the nine standards as conducted by external KDE audit teams) provided a means for Kentucky schools to measure their ability to increase school achievement. As a result of Scholastic Audits, it was determined that several Kentucky schools were in need of assistance to improve their educational programs.

The Comprehensive School Reform Demonstration Program (CSRD) provided assistance to schools that were in need of improving their educational programs and student achievement. Many Kentucky schools used Title I funds to implement various school reform models including Accelerated Schools Project, America's Choice, ATLAS Communities, Early Intervention in Reading, First Steps, Modern Red Schoolhouse, School Development Program, Success for All (St. John et al., 2000). The results of the Scholastic Audit determined the school’s area of focus which guided the selection of the reform model to be implemented.

Nine Kentucky elementary schools selected America’s Choice comprehensive reform to address their school improvement needs based on their Scholastic Audit. This quantitative study--based on policy linking the educational history of school reform, comprehensive design models, America’s Choice, and discipline-based achievement--examines mathematics outcomes in the seven schools which had data for math. Specifically the central research question summarizing this study is: What are the effects of implementing America’s Choice comprehensive school reform design in mathematics in seven elementary schools in Kentucky?
CHAPTER II

REVIEW OF LITERATURE

Introduction

For many years, American presidents and the federal government have been interested in improving the performance of students who come from low-income homes. The work and support of past presidents led to the development of Lyndon Johnson’s “War on Poverty” initiatives in the 1960s. The result of this initiative led Congress to enact the Elementary and Secondary Education Act (ESEA) 1965. The ESEA was a federal education bill authorizing funds for professional development, instructional materials, resources to support educational programs, and promotion of parental involvement. The lack of student improvement on state mandated achievement tests caused legislatures to be concerned about the long standing achievement problems in public schools. Because of inertia related to achievement—both overall levels and continuing inequalities—the government has reauthorized numerous versions of the ESEA since it was first enacted. The current version of ESEA is the No Child Left Behind Act of 2001.

Over the past several decades, a sense of urgency has propelled educators to search for and implement effective strategies to improve schools across America. The report, *A Nation at Risk* (National Commission on Excellence in Education, 1983), focused attention on the academic achievement levels of students in America compared to that of students in other countries. Legislatures and policy makers were concerned with the United States’
low ranking and decided to take action by calling for comprehensive reform in low performing schools (Anderson, 2002; Berliner & Biddle, 1995; Bracey, 2004; Lasley, Siedentop, & Yinger, 2006). Nationwide, educators responded to the call for comprehensive reform as the key to lasting school improvement for low performing schools.

In the past, school reform initiatives followed a cycle where high level policy talks developed into reform programs. However, these programs focused on one aspect of reform and typically showed little impact on achievement, leading to the next reform cycle. Comprehensive School Reform (CSR) represents an attempt to break away from this pattern in order to meet the educational challenges of today. CSR programs encompass all aspects of school operations including instruction, assessment, classroom management, professional development, parental involvement, school management, and curriculum (Ross & Gill, 2004; Rowan, Camburn, & Barnes, 2004, Schmoker, 1996).

This study examines the impact of implementing the America’s Choice comprehensive school reform design for seven Kentucky elementary schools in the area of mathematics achievement for third grade exiting primary students to determine the impact of this program, specifically scores for the years prior to implementation, the years during supported implementation of the America’s Choice comprehensive reform design, and the year after support ended.

In this chapter the researcher presents literature reviewed, organized into sections relevant to the conceptual framework of the current study. The first section relates to school reform generally; the second to comprehensive school reform. Next, the author examines one specific approach to comprehensive school reform, America’s Choice. Then, mathematics in the United States is addressed. The chapter concludes with a
summary connecting the literature to the research problem and purpose for this study.

School Reform

The past decade has seen most school districts in the United States engage in some type of school reform movement either by choice, mandate, or both. Educators live in a time of numerous opportunities for improving the education of students and the school environment. The No Child Left Behind Act of 2001 (NCLB) shaped the work of public school teachers and administrators in the United States by incorporating standards-based accountability (SBA) provisions which provide for an increased emphasis on student achievement (Stapleman, 2000). The components of a standards-based accountability system include aligning standards and assessments, rating school performance, reporting performance, providing assistance, and creating consequences (Craig et al., 2005; Desimone, 2002).

Until a few years ago, most reform efforts focused on individual subject areas like mathematics, reading, writing, or science rather than looking across all curriculum areas. McCombs (2002) has suggested that effective school reform models include the following eleven basic components: (a) research based methods, (b) comprehensive design, (c) professional development for staff, (d) measurable goals and benchmarks for content areas, (e) staff support within schools, (f) support for teachers and principal, (g) parental and community involvement, (h) external technical support and assistance, (i) annual evaluation, (j) coordination of resources, and (k) strategies to improve academic achievement.

Researchers suggest that aligning organizational structure and instructional practices with common goals to produce an effective school organization. To this end, schools with well-integrated, coordinated approaches to teaching and learning generally
have higher student achievement. A school-wide approach to school reform may be more effective in producing gains in student achievement than separate initiatives that target individual aspects of the school (Michael & Young, 2005).

In addition to the basic components of effective reform models, Dimmock (2002) shares some common aspects of school reform that include the following concepts: (a) origin of design, (b) goals of the school improvement intervention, (c) targets of change in the schooling process, (d) specification of the design, (e) opportunity to learn by teachers, (f) connectivity-consistency, (g) implementation process, (h) leadership-participation and control, (i) outcomes and effects, and (j) match between design model and school context. Schools that adopt school wide reform initiatives must evaluate and address these concepts in order to develop a strong schoolwide reform model. These concepts allow schools to look closely at different reform models to determine the one that best meets their needs in order to be the most effective in producing the desired results.

School reforms are typically either curriculum-based models centered on a particular content area or process-oriented models designed to improve the educational outcomes of students. Some curriculum based reforms focus on one academic area but most recent models include multiple content areas. There are a number of proven reform models available with strategies ranging from targeted interventions for students who are having difficulty learning to read in the regular classroom (e.g., Direct Instruction and Success for All--Ross, Nunnery, Goldfeder, Rachor, Hornbeck, & Fleishman, 2004) to comprehensive reforms introducing structured curriculum to the school (Dimmock, 2002; Mason, Mason, Mendez, Nelson, & Orwig, 2005; McCaslin & Good, 2008). Each type of reform model has its own strengths and limitations; accordingly it is crucial for educators
to evaluate closely each one relative to context before choosing the appropriate model.

Some school reform models are curriculum-based; these are typically implemented to assist schools who have identified one curriculum area as their weakness. To determine the effectiveness of such focused disciplinary-based reform models, researchers have studied reforms that focus on reading, math or writing. For example, Peterson, McCarthey, and Elmore (1996) conducted a longitudinal analysis of three restructuring experiments in three elementary schools located in large urban school districts across the United States that were experimenting with restructuring writing content. The research question addressed the nature of restructuring in relation to its effects on instructional practice in writing.

Peterson et al. (1996) used purposeful selection to find three elementary schools that were undertaking school-wide restructuring. Six teachers (two from each school) who taught writing participated in the case studies. School 1 is an urban school in the northwestern United States with an ethnically diverse group of 245 students. The student population consisted of 55% minority with 30% qualifying for free and reduced lunch. School 2 is located in a major metropolitan area in the central southern United States with a student population of approximately 500 students. Students are bused to this school from other areas of the town which accounts for the 26% African American students and 25% of the students qualifying for the free or reduced lunch program. School 3 is located in a poor neighborhood in a densely populated urban area in the eastern United States. The population consisted of 250 students from ethnically diverse backgrounds (90% minority). About 80% of the students qualify for the free or reduced lunch program.

Data collection consisted of full-day classroom observations in which teachers wore wireless microphones supplemented by dish microphones which provided data on
student-to-student interactions. Samples of student writing were also collected as part of
the data analysis. Researchers also attended one staff meeting per school as well as other
school related events in order to gain information about the school culture. The
researchers developed two classroom protocols to gain information from classroom
teachers: (a) new roles (focused on decision making, collegiality and collaboration, and
conceptions), and (b) post-observation formats. Final data were collected from post-
observation interviews to clarify information (Peterson et al., 1996). Data analysis
consisted of verbatim transcription of all interviews and observations using analytic
induction to look for emerging patterns. Triangulation occurred through the use of several
sources of data and continual comparisons. Emerging patterns include (a) physical
workplace, (b) collegiality, (c) teacher roles within the school, (d) decision making, and
(e) opportunities for professional development. The emerging categories were used to
compare the three schools in the case study.

Peterson et al. (1996) found that all three schools successfully restructured
according to each school's vision of restructuring. The reorganization at each school
shared four key features: (a) all three schools had a philosophy relating to student
learning, (b) teachers met on a regular basis to discuss curriculum issues and instruction,
(c) teachers at all schools were involved with shared decision making, and (d) teachers
had access to new ideas about instruction through staff development or ongoing
discussions about teaching.

When schools evaluate their data and determine a school wide need to restructure
the school, staff can effectively select and implement a program that will guide them
toward making the necessary alterations in program implementation that lead to
successful change. The information about restructuring provided from these case studies
would be beneficial to administrators and teachers from schools of similar backgrounds considering restructuring in their district or buildings. In order to implement a restructuring model in a school effectively, it is helpful to examine the impact of implementation on schools that are demographically similar (Peterson et al., 1996).

In 2005, Michael and Young conducted a case study to capture the prominent themes and educational practices at the building level that administrators perceive as best educational practices. The study had a two-fold purpose: (a) to gain an understanding of how senior school administrators define inspired public schools, and (b) to discern the characteristics of inspired schools to guide meaningful school improvement efforts. The researchers use the term inspired schools to refer to a school whose staff shares common values and intentional practices. This term is an extension of effective schools that focus primarily on exemplary student results and resilient schools that deliberately teach students how to thrive despite challenges and adversity in life. Participants were randomly selected from the New England area based on fit with the qualifications of being an inspired school. Eight schools and twenty-nine ($N = 29$) senior administrators (school superintendents and assistant superintendents) were randomly selected to participate in the study. Nineteen senior administrators opted to participate ($n = 19$). Each school represented different demographics, levels, and district sizes.

Data collection consisted of conducting in-depth interviews with administrators and building personnel. Responses were transcribed and categorized into themes. Eleven major themes emerged from the data collected (Michael & Young, 2005). Theme-oriented qualitative data analysis techniques were used by Michael and Young to identify the following eleven themes: (a) attention to developmental needs of all members of the school community, (b) pervasive leadership, (c) a relational approach to education, (d) a
deliberate "assets-based" approach to student learning to promote efficacy and resiliency, (e) traditions that nurture a sense of belonging and mark rites of passage, (f) a unique sense of place and mission, (g) reliance on an inner compass for the school knowing the direction and steps toward reaching the goals, (h) intentionality in promoting a sense of inclusively, equity and global citizenship, (i) celebration of large and small "victories," (j) a high sense of stakeholdership in the school, (k) a commitment to community involvement and service, and (l) a visible valuing of the integrated arts. Findings indicate that the development of a set of common characteristics could be identified for school administrators looking to promote meaningful school improvement at the building level. Limitations include the fact that only a small number of administrators and districts chose to participate in the study.

McCombs (2002) evaluated a school reform model, the Community for Learning (CFL) program, to illustrate using the American Psychological Association's (APA) Learner-Centered Psychological Principles, The Learner-Centered Framework to assess the effectiveness of comprehensive school reform models. The research focused on determining if the school's program is learner centered. Participants were three east coast inner city schools implementing the CFL program. The sample size included seven fourth-grade classes and six fifth-grade classes for a total of 13 teachers and 286 students.

Data collection consisted of teacher and student surveys and the Degree of Implementation Assessment Battery for Adaptive Instruction developed by Wang (cited in McCombs, 2002). This instrument is a systematic measure that examined teacher's performance on 12 critical dimensions. The mean implementation percentages across all categories and teachers were ranked from lowest to highest and the midpoint was selected. The means were compared to determine significant differences between high
and low CFL implementers. Data analysis showed that students of teachers who were high CFL implementers scored significantly better (an average of 3 to 4 points or 8%) than students of teachers who were ranked as low CFL implementers in the areas of reading and math. Significant correlations were found between classroom achievement scores and standardized reading \((r = .38)\) and math \((r = .46)\). The findings from the McCombs study demonstrate and confirm the need to select a comprehensive school reform model that focuses on both research-based practices and learner-centered approaches to instruction. Schools can focus their instruction on a learner-centered approach without selecting an approach that is research-based. This study demonstrated the need and results of selecting a model that is both in order to see positive results in the classroom and school.

It is important to review and select a school reform model that has research-based practices and learner-centered approaches to instruction. Sometimes this focus is in one specific content area. The school data evaluation provides a guide to the administrators in identifying and selecting a single-subject focused school reform model or a school-wide reform model. Dimmock (2002) conducted a case study analysis to apply a developed framework as a means of describing and analyzing a particular well-developed school design model. The school design model is a type of reform that looks at the school from a systemic point of view as opposed to focusing on just one content area. All aspects of the school are addressed in the reform model. The study focused on the implementation of the school design program and the role played by the school leader and change agent. The investigation presented a classifying framework of 10 criteria aimed at capturing the characteristics of school design programs and their leadership and implementation.

Purposeful sampling was used to select the participating school which was located
in a rural part of New England, based on the representation of the phenomenon to be studied. The small secondary school serves a middle and upper socio-economic group with 350 students coming from other parts of the United States and other countries (Dimmock, 2002). Data collection centered on the ten criteria: (a) origin of design model selected, (b) goals of the school improvement intervention, (c) targets of change in the schooling process, (d) specification of the design, (e) opportunity to learn by teachers, (f) connectivity-consistency, (g) implementation process, (h) leadership-participation and control, (i) outcomes and effects, and (j) match between design model and school context.

Data were collected over a three-year period consisting of interviews with administrators and teachers, classroom observations, and standardized test scores. Results showed a positive impact from the implementation of a school design model. Significant improvement was found in scores across all core academic subjects, as well as an increase in positive student behavior and a decrease in discipline issues—all improving as the level of implementation increased. Dimmock (2002) found the framework developed and studied to be a useful guide in evaluating the effectiveness of a school design model. Implications from the case study led future researchers, both qualitative and quantitative, to reevaluate the classification of school reform programs of the future. When schools are reviewing and selecting an appropriate reform model to meet their needs, it is important to understand how to evaluate the individual frameworks of each reform model.

Summarizing this section, Peterson et al. (1996) and Michael and Young (2005) studied the nature of school restructuring from the administrative point of view and the teachers’ point of view to determine the outcome and impact on the overall school environment. Both studies provided information to show that the implementation of school reform is complex to implement correctly but does produce positive outcomes.
McCombs (2002) and Dimmock (2002) also added to the research in this area. Both of these studies developed and tested methods to assess the effectiveness of various school reform models. McCombs focused on the fact that school reform must be research-based and student-centered to achieve the desired outcomes. Dimmock’s research evaluated a useful guide to evaluate the effectiveness of the school reform model implemented. As schools work to implement school reform models to address learning needs, they must have in place an effective method of evaluation for the reform model.

Various models of school reform exist to meet the individual needs of the schools. Schools must first complete a systematic evaluation of all available data to make an informed decision as to what the main areas of focus should be to enable them to select the appropriate school reform model to best meet their educational needs. Once it is determined that a school is in need of assistance, the leadership must select the correct school reform model to implement in order to achieve the desired results (Michael & Young, 2005). The goal of school reform is to provide a path for schools to increase student achievement. Selecting the appropriate school reform model is essential to this quest.

Comprehensive School Reform

The goal of school reform is to produce real learning sustained over time and not just an initial increase in standardized test scores. For this to occur the reform must be comprehensive. Many believe that Comprehensive School Reform (CSR) holds real promise for improving schools because it aligns all parts of the system—standards, curriculum, instruction, assessment, classroom management, professional development, parental involvement, and school management—which leads to student success.

In overall low performing schools, the needs are so pervasive that a whole system
approach, such as Comprehensive School Reform (CSR), is often more efficacious. Recent studies reporting evidence of the effects of CSR on academic achievement have found most of the results to be favorable. Because Comprehensive School Reform programs incorporate all aspects of school operations, the systemic change has a greater likelihood of impact on student achievement.

Over the past two decades, Comprehensive School Reform (CSR) has evolved from theory into widespread implementation. This evolution is in part due to the report *A Nation at Risk* (National Commission on Excellence in Education, 1983), which stated to the public that individuals leaving schools and entering the work force in America were unprepared and unable to compete internationally. This caused the public to question the educational system. Schools moved to address these concerns and prepare the nation’s youth to enter and compete in society more successfully.

The Obey-Porter Comprehensive School Reform Demonstration (CSRD) (1998) legislation provided a monetary incentive program to provide the opportunity for hundreds of schools across the nation to try comprehensive school reform models. In 1999-2000, the CSRD funded more than 1,800 schools nationally with a minimum $50,000 each for three years (Hatch, 2001). By 2002, over 380 “models” were adopted with CSRD support (Desimone, 2002).

The commitment by Congress and the United States Department of Education to improve schools through comprehensive school reform is underscored by CSR’s incorporation into the No Child Left Behind Act of 2001 (NCLB) (Pub. L. No. 107-110, 2002) as a formula grant with allocation based on the district’s Title I allocations. Comprehensive School Reform Demonstration (CSRD) grants provide three-year funding to low-performing schools through competitive federal grants. Comprehensive school
reform (CSR) focuses on reorganizing and revitalizing entire schools, rather than implementing a number of specialized, and potentially uncoordinated, school improvement initiatives (Borman, Hewes, Overman, & Brown, 2003).

Throughout the last decade many policies have been enacted with the intention of raising overall achievement for disadvantaged students (Tyack & Cuban, 1995; Hurst, Tan, Meek, & Sellers, 2003). These policies included academic standards, targeted resource programs such as Title I, decentralized initiatives like the restructuring movement, site-based management, and performance assessment systems, e.g., those mandated by the 1994 reauthorization of the Elementary and Secondary Act. Many researchers have conducted studies to evaluate and provide information on the various CSRD models that exist. School administrators using this information make informed decisions regarding the selection of the best program for effective implementation in their school.

Beam and Faddis (2001) reported findings from the Comprehensive School Reform (CSR) Implementation survey conducted by the North Central Regional Educational Laboratory (NCREL). This study assessed the extent to which schools receiving CSR funding focused on developing each of the nine comprehensive school reform program components. The components included (a) comprehensive school reform design, (b) support within the school, (c) measurable goals and benchmarks for student performance, (d) parental and community involvement, (e) effective research-based methods and strategies, (f) professional development, (g) coordination with the schools other reform efforts, and (h) evaluation strategies. Participants were teachers and students in CSR schools from Illinois, Indiana, Iowa, Michigan, Minnesota, Ohio, and Wisconsin. Surveys were sent to 361 schools ($N = 361$) with a response from either teachers or
principals from 221 schools \((n = 221)\). This provided an overall return rate of 61% with principals at 58% and teachers 54%. Schools who returned surveys were compared to schools that did not return surveys to check for overall representativeness of the sample group in the following dimensions: (a) geographic locale, (b) grade level, and (c) Title I status. A chi-square test revealed urban schools were less represented in the returned sample, \(\chi^2(2, n = 345) = 8.17, p < .05\).

Data collection consisted of principal and teacher responses to surveys evaluating the primary and secondary foci of the implementation process and their progress in achieving the goals. A four-point Likert-type scale was used to evaluate the survey responses. Data analysis identified the following six areas of primary focus; (a) develop a comprehensive reform plan, (b) obtain staff support, (c) establish measurable goals and benchmarks, (d) use effective, research-based strategies to improve curriculum, (e) use effective, research-based strategies to improve instruction, and (f) provide professional development. Beam and Faddis (2001) reported the following results from the data analysis. Principals and teachers demonstrated moderately high ratings of progress on most reform components. The principals’ responses were developing comprehensive plan \((M = 3.77)\), attaining support staff \((M = 3.78)\) and providing professional development \((M = 3.75)\). The teacher responses recorded developing comprehensive plan \((M = 3.67)\), attaining support staff \((M = 3.47)\) and providing professional development \((M = 3.57)\).

From the results it is evident that the principals’ ratings of progress were significantly higher than teachers’ ratings for 10 of the 15 components. Overall, survey findings revealed a high level of agreement among principals and teachers with respect to the degree of focus of the reform effort and the progress made toward accomplishing the goals. The information gathered by Beam and Faddis (2001) would be beneficial to
schools or districts considering implementing some form of school reform in their building or district. Once schools determine their areas of weakness they can use this information to guide their selection of an appropriate school reform model that would meet their needs.

Holdzkom (2002) studied the degree that Comprehensive School Reforms (CSR) brought about change in student outcomes, teachers’ practice, and teachers’ perceptions that characterize the school. The researcher looked at (a) what differences in classroom practices, school climate, and reading achievement exist between CSR and control schools after the effects of location (urban and rural) had been controlled, and (b) what differences exist among CSR models in terms of classroom practices, school climate, and student achievement after at least three years of implementation. Participants represented twelve schools from Kentucky, Tennessee, Virginia, and West Virginia. Each school adopted one of the following comprehensive school reform models (a) Different Ways of Knowing (DWoK), (b) Success For All (SFA), (c) Core Knowledge (CK), (d) Direct Instruction (DI), or (e) Balanced Early Literacy Initiative (BELI). Each study school was matched with a control school which shared similar demographic characteristics.

Data collection consisted of teacher observations, surveys, teacher and staff interviews, and instruments that tested student’s reading abilities. The School Climate Inventory (SCI), an instrument developed at CREP in 1989 (cited in Holdzkom, 2002) was used to monitor and manage the data. Findings showed (a) achievement gains in the third year of implementation in CSR schools was significantly higher than in non-CSR schools, (b) Faculty and staff at CSR schools appeared more likely to be positive in perceptions of order, instruction, and expectations, (c) emerging differences in instructional practices appeared related to the selected comprehensive school reform
model, (d) different models evidenced different aspects of instruction, classroom practice, and teacher-student relationships. The findings from the study led administrators, school district personnel, and building administrators to evaluate not just the reform model but to look more closely at the strategies addressed within each model. The No Child Left Behind legislation opened the door for more research in determining best strategies to help educators meet this goal.

Sterbinsky, Ross, and Redfield (2003) conducted a longitudinal study to examine the effectiveness of Comprehensive School Reform (CSR). Three questions guided their study: (a) Is there a difference in classroom practices between CSR and control schools after two years of implementation and do the differences vary by location, (b) Is there a difference in school climate between CSR and control schools after two years of implementation and do the differences vary by location, and (c) Is there a difference in reading achievement between CSR and control schools after two years of implementation and do the differences vary by school location?

The participants in the study were twenty-four schools in Kentucky, Tennessee, Virginia and West Virginia. Twelve schools were CSR schools and twelve were control schools that matched the CSR schools on the basis of free and reduced lunch, rural or urban, number of students and grade levels, and results on state mandated tests. Sterbinsky et al. (2003) visited each school multiple times to conduct classroom observations, implement teacher surveys, and administer reading tests to the students. Over a two-year time span, data were collected from 3,500 classroom observations, 1,100 teacher surveys and 1,300 student reading tests. The researchers used MANOVA to analyze the observations and school climate and MANCOVA for the reading achievement data.
Classroom observations had specified classroom practices for the dependent variable and CSR status, location, and year for the independent variables. School climate was analyzed with school climate dimensions as the dependent variable and CSR status, location, and year of implementation for the independent variables. Reading achievement used 2001 reading achievement scores as the dependent variable with CSR status, location, and year of implementation for the independent variables (Sterbinsky et al., 2003). The results indicated that CSR teachers used direct instruction and performance assessments. CSR teachers were significantly more positive about leadership, support, capacity, pedagogy and outcomes. Implementing a CSR Model also impacted reading achievement with CSR schools out-performing control schools on two reading assessments. In interpreting the findings of this study one must consider if the schools in this study are representative of all schools that implement CSR Models. If this is true, then these findings would be generalizable to other schools that are similar. Policymakers and district-level educators of schools implementing a CSR Model would find this study useful in their decision making process.

Good, Burross, and McCaslin (2005) conducted a longitudinal study of 24 schools in Arizona that received CSR funding over a six-year period. Participants consisted of 24 schools implementing CSR grant programs and 24 matched schools as the control group in Arizona. Schools were matched on the basis of geographic location, enrollment, classroom expenditures per pupil, and percent of students receiving free or reduced lunch. The study compared the achievement gains of third through ninth graders in CSR and non-CSR schools over the six year period. The study focused on two issues: (a) student longitudinal grade-level performance, and (b) a cross-sectional analysis comparing fourth graders’ performance over a two-year period. The purpose of the study was to determine
if students became more competent the longer they were in a school implementing the CSR program. Student performance was analyzed to determine if student achievement increases the longer they are in CSR schools.

Data from Stanford-9 reading, language, and mathematics assessments from 1996-97 through 2001-02 were collected and analyzed from each of the participating schools. The 24 CSR schools provided a total of 1,536 performance area scores and the non-CSR schools produced 1,614 performance area scores. National percentiles were converted to normal curve equivalent (NCE) scores to perform parametric statistical analyses based on an equal interval scale. Mixed design MANCOVA tests were performed on scores by performance areas (reading, language, and math). CSR and non-CSR schools showed similar growth patterns in reading, math, and language. Paired sample tests on the means showed no mean difference in any performance area [reading, \( t(75) = 1.06, p = .29, \eta^2 = .01 \); language, \( t(75) = .43, p = .67, \eta^2 < .01 \); math, \( t(75) = 1.19, p = .24, \eta^2 = .02 \)]. Findings indicated that nontrivial gains in performance (1.0 standard deviation in math, .62 in reading, and .11 in language) were found in the combined samples of CSR and non-CSR schools. From these findings one could conclude that the schools in this study implemented various aspects of comprehensive school reform models in order to make the gains needed for achieving accountability goals (Good et al., 2005).

In order to achieve consistent and sustainable growth in achievement schools must continue to implement all components of the CSR model effectively. Each model was designed with certain components that work together to achieve a systemic change within the school. Epstein (2005) conducted a longitudinal case study to examine the implementation and results of the Partnership Schools Comprehensive School Reform
model in a Title I elementary school. Two guiding research questions were (a) can the model be implemented and (b) if so, what are the school results on the state standardized achievement tests. The participating site for the case study is a Title I school with a student population of 375 students in grades K through 5 located in an urban district in Connecticut. Of the student population 51% participate in the free or reduced lunch program and 40% are English Language Learners. The school received a grant to implement the Partnership Schools Comprehensive School Reform (CSR) model to improve school, family, and community partnerships and student achievement.

Epstein (2005) collected data from various sources including (a) detailed four-page action plans, (b) reflective five-page end-of-the-year evaluations from each action team, (c) school artifacts, (d) end-of-the-year school update surveys on program development, (e) two site visits each year, and (f) student achievement test scores. The study findings indicated that the CSR School implemented structures and processes for the five policy attributes (specificity, consistency, authority, power, and stability) that affect the quality of CSR programs. Longitudinal data revealed that the student scores on state achievement tests in mathematics, reading, and writing showed significant improvement over the three years of the study. This study would be of interest to administrators in districts with low performing Title I schools who are looking to develop a plan to increase school achievement because the schools in the study that implemented quality CSR programs showed significant improvement on state achievement tests.

Wetherill and Applefield (2005) conducted a qualitative study to examine and explain the wide discrepancy in outcomes among schools implementing school reform models. Eight North Carolina schools in their first year of the federally funded Comprehensive School Reform Development (CSRD) program participated in the study.
Four schools achieved positive project initiation and four schools experienced a less positive impact. The schools selected represented urban and rural elementary, middle, and high schools and came from all regions of the state. Student populations differed in number, racial make-up and free and reduced lunch qualifications. Data collection occurred from fall and spring on-site visits to each school. The visits included principal interviews and teacher group interviews. Interviewer notes and taped responses were analyzed for evidence of the four states of change: (a) State I-Premature, (b) State II-Hesitant, (c) State III-Developing, and (d) State IV-Established. Data from interview responses were correlated with observation notes to establish profiles for each school.

Data analyses of the school profiles categorized each school into one of the change states. Three schools were in State I (Premature), one school was State II (Hesitant), four schools were State III (Developing), and no schools were State IV (Established). The profiles gave insight into the effectiveness or lack of effectiveness in the implementation of the school reform model. Schools that had little support or low levels of implementation were in State I while schools that had higher levels of support and implementation were at State III. Wetherill and Applefield (2005) found that planning and implementation of the reform model must include teacher’s involvement in the collaborative efforts and must be responsive to the existing characteristics of the school environment. Administrative leadership is essential in determining the change state of a school in order to lead the school to the next level of change implementation. An important implication of this study is that leadership strategies should vary depending on the change state of the school. Research shows that school or administrative leadership has proven to be a key element in the successful implementation of any reform model. A limitation to this study is the fact that student outcomes were not addressed or connected.
with the state of change at any of the schools.

In sum, Beam and Faddis (2001), Epstein (2005), and Holdzkom (2002) studied the impact of implementing comprehensive school reform program components. They found that to achieve the desired outcomes, all of the components must be implemented. Good, Burross, and McCaslin (2005) conducted a longitudinal study to determine the impact on students over time and found that effectively implementing all aspects of a CSR program, when sustained, had an overall impact on student outcomes. In looking at the various program components, the role of the school leadership was found to be important in the implementation of any CSR model (Wetherill and Applefield, 2005; Sterbinsky, Ross, and Redfield, 2003).

In implementing any model of comprehensive school reform it is important to gain administrative support to achieve the desired outcomes. This is true to have successful implementation of any instructional program. Effective classroom instruction should be maintained in all schools to reach and maintain levels of student achievement. McCaslin et al. (2006) completed an observational study to examine literacy and mathematics instruction in 20 low-income schools enrolled in CSRD programs in Arizona. The goal of their research was to determine if instruction in schools participating in CSR programs differed from normative instruction in important ways. They collected research data from 20 elementary or K-8 schools throughout Arizona. The student demographics in the schools included poverty rate (defined as the percentage of students receiving free or reduced lunch), ranging from 41% to 100% with a mean of 85% and student mobility ranging from 12% to 71% with a median of 36%. The sample included both small and large schools within rural and urban contexts. The student populations were very diverse and included African American, Hispanic, Native
American, and white students.

McCaslin et al. (2006) conducted reading and mathematics classroom observations over three semesters: spring, 2003, fall, 2003, and spring, 2004. Third grade observations included 54 teachers and 133 classroom visits; fourth grade, 52 teachers and 173 visits; fifth grade, 39 teachers and 141 visits. The total for their research included 145 teachers during 447 visits. The observations can be summarized as 587 ten minute intervals of mathematics instruction and 1,642 ten minute intervals of literacy instruction. The ten minute cycle included five minutes of narrative for classroom events, activities and affect; two minutes for rest and reflection; and three minutes for coding the observed events. Data gathered and reported from the observations focused on instructional opportunities, student activities, and the nature of the teacher and student relationships. The variables for the research represented 19 categories and were coded as present or absent. The observers received training and practice so their work and recordings would be reliable for this project. The data analysis showed that normative practices in the 20 CSR schools involved curricula, instruction, and management that are primarily and coherently focused on basic facts and skills related to elaboration and thinking. The normative instruction is highly similar to normative instruction in grades one through five nationally.

Administrators and teachers should develop a plan to implement the components of the CSR model effectively in order to achieve school change thus impacting student achievement. Schools that are able to accomplish this task successfully will see a rise in student achievement scores. There are various CSR programs available for schools to select. Administrators need to become familiar with and analyze each to determine the most effective model for their school.
Ross et al. (2004) examined the effectiveness of two of the most widely used Comprehensive School Reform programs in an urban school district in Ohio. The specific research questions were: (a) what were the impacts of Direct Instruction (DI) and Success for All (SFA) on student achievement outcomes over time, (b) what were the program impacts on school climate, and (c) what school and district variables influenced the effectiveness of the school's implementation of the program? Two groups participated in the three-year study, one consisting of six Direct Instruction (DI) schools and two Success for All (SFA) schools and a matched control set of schools. The DI schools served predominantly African-American, high poverty student populations. The SFA schools' student population was similar except for lower poverty rates. The control group was matched based on socioeconomic factors (SES), prior achievement, and ethnicity.

Student-level and school-level analyses were conducted using a two-level hierarchical linear model (HLM). Data was collected from Stanford 9 Total Reading in grade two and Ohio Proficiency Tests in grades four and six. Pretest and posttest means were analyzed to determine program effectiveness (Ross et al., 2004). Schools implementing the DI program had significantly lower posttest means in fourth grade \( (t = -2.37, df = 40, p < .05) \) and sixth grade \( (t = -2.30, df = 40, p < .05) \). There were no significant differences between DI and other schools after adjusting for poverty rates. Overall results for each model failed to reveal significantly different pretest-posttest scores for district schools with similar demographics. Individual school achievement test results for both models were mixed.

Findings of the Ross et al. (2004) study would be beneficial to policy makers and administrators responsible for the selection and implementation of Comprehensive School Reform models. It is also important to consider not only the outcomes based on
the number of years of implementation of each program but also differential demographic context. A limitation of this study is the fact that the schools were in their second year of implementation of their programs. In selecting and implementing a CSR program, administrators should consider school demographics when examining the success rate for different CSR models. How model effectiveness is defined and varying degrees of success for students from different backgrounds impacted decisions made about selecting a model to implement.

Once administrators evaluate and select the appropriate CSR model to implement in their school, a mechanism to evaluate the progress and impact of the implementation of the program should be developed. Carlson (2003) used a case study design to conduct a follow-up study of five small isolated southwestern schools to determine if the schools were able to sustain their initial efforts in implementing their CSR models, and if the schools observed any difference in student performance. Two general questions guided the study: (a) how successful was each rural school in fully implementing their CSR model, and (b) what differences emerged in student performance over the three years of the grant. Participating schools were selected based on their participation in the first study after completion of year one of implementation of their reform models. The schools were selected to provide a cross section of characteristics including size of school, grade levels, size of community, location respective to metropolitan areas, and CSR model chosen. The schools represented the states of Louisiana, Arkansas, New Mexico, Oklahoma, and Texas. The CSR models implemented included Core Knowledge, Early Literacy Initiative, Success for All, Effective Schools Models, and Reading Renaissance.

Data collection for the study included on-site observations and interviews with administrators and teachers and telephone interviews with model developers and
consultants. Data from pertinent reports and test results were analyzed to determine student success rates for each school (Carlson, 2003). The study showed that schools selecting more prescriptive reform models implemented their programs more quickly and achieved student success at a faster rate than the schools who selected a less prescriptive reform model. In general all five schools were reasonably successful in implementing and sustaining their respective CSR plans. Each school represented a different state and a different model and therefore the researcher was unable to make school to school comparisons of reading and math achievement scores. The test results for each of the schools in this study provide an inconclusive picture of the student success rate for the different CSR models.

Selecting and determining the success of a CSR plan depends on (a) financial support, (b) classroom practice change, and (c) support or use of scientifically based educational interventions. CSR efforts are essential to stimulate new instructional efforts that have the potential to make a difference in learning in any school setting. Selecting a descriptive CSR model provides the needed support for schools to implement the program effectively and move towards the desired outcomes.

In order to maintain the growth in achievement achieved during the initial years of implementation of a CSR model, it is important to lay a good foundation for the teacher staff in relation to effective implementation of all components with internal support only. It is much easier to implement a new program with support generated both internally and externally. The goal of implementation is to learn to maintain the level of implementation from internal support only to continue the growth in student achievement.

Sterbinsky et al. (2005) conducted a three year (1999-2001) quasi-experimental
(matched treatment-control groups) study of CSR implementation and outcomes in multiple schools. In the four state area of Kentucky, Tennessee, Virginia, and West Virginia, 140 schools received funding to implement CSR programs. For this study, 12 CSR schools were selected and matched with 12 non-CSR schools that matched them demographically for participation in this study. The CSR models that were implemented in the schools and examined in the study included Success for All (SFA), Different Ways of knowing (DWOK), Balanced Literacy Initiatives (BELI), Direct Instruction (DI), and Core Knowledge (CK). The two broad research questions that guided the research and analyses were (a) what differences in classroom practices, school climate and reading achievement occurred between CSR and control schools over a three-year period and (b) did CSR impacts vary for urban and rural schools?

Data analyses included classroom observations to capture data related to instructional orientation and strategies, classroom organization, student activities, technology use and assessments. Teacher surveys were collected to analyze information on school climate, teacher expectations and perceptions on outcomes. Student achievement was assessed using individually administered standardized reading tests. Sterbinsky et al. (2005) used three way MANOVAs (Program, location, and year) for the primary statistical analysis design for most dependent variables. Effect sizes (ES) were computed using Cohen’s $d$ formula where appropriate. During the three years a total of 413 (128 in 1999-2001, 146 in 2000-2001, and 139 in 2001-2002) School Observation Measures (SOMs) were conducted. During the analyses all three main effects were significant: Year [$F(52, 568) = 4.47, p < 0.001$], Program [$F(26, 285) = 6.72, p < 0.001$] and Location [$F(26, 285) = 9.11, p < 0.001$]. All three multivariate interaction effects were also significant: Year x CSR [$F(52, 568) = 2.85, p = 0.001$], Year x Location [$F(52,
568) = 2.8, p < 0.001] and CSR x Location \( F(26, 285) = 2.68, p < 0.001 \) and the three way interaction \( F(52, 568) = 1.96, p < 0.001 \).

These findings represent only a snapshot of what actually occurs daily in these classrooms. They do not imply that the models were necessarily well implemented in each of the schools throughout the entire day. The results from within and across the three sets of analyses (longitudinal cohort, gain score, and yearly cohort) were favorable for all CSR schools. CSR students had significantly higher gain scores than the control students demonstrating strong effect sizes ranging from +0.31 to +0.44 (Sterbinsky et al. 2005).

Zhang, Fashola, Shkolink, and Boyle (2006) conducted a study to examine the implementation of CSR and changes in reading and math school level achievement data in 17 states. The researcher did not list the cited states. The focus was to improve understanding of the impact of implementation of CSR models on student achievement. For the purpose of this study, they focused on three hypotheses. First, changes in implementation level and student achievement are sequential and may be a function of implementation year. Second, the number of years of implementation and the level of implementation are not necessarily correlated over time. Third, the specific level of implementation does not necessarily correspond to the same amount of improvement in student achievement among different CSR models implemented by schools.

Participating schools selected for this study were part of the larger National Longitudinal Evaluation of Comprehensive School Reform (NLECSR) study consisting of 649 schools from 21 school districts in 17 states. A propensity score approach was used to match CSR schools with non-CSR schools within each of the districts in each state. From this process 138 pairs were generated. Zhang et al. (2006) selected 115 pairs
with school-level achievement data available to represent their sample for the study. The researchers gathered survey data and achievement data. Survey data addressing the core components of CSR models (school leadership, professional development, and pedagogy) and demographical information were collected from principals and districts during the 2002, 2003, and 2004 school years. The achievement data collected consisted of five years (1999 to 2003) of school level achievement data. The data collected included two years before the distribution of the survey and three years after the survey was administered.

Data analysis consisted of a two level (year at Level 1 and schools at Level 2) time-series approach in hierarchical linear modeling (HLM) to model the improvement in student achievement. The analysis showed that adopting a specific CSR model did not guarantee any effect on student achievement. The implementation of the CSR model is the key to determining any impact on student achievement. Zhang et al. (2006) found that the impact of CSR implementation on student achievement is conditional on several factors including number of years of implementation, implementation level, and the specific CSR model being implemented.

Many researchers have focused their research on which types of CSR models produced the best or greatest impact on student achievement. This input-output approach provides limited data to assist administrators in making evaluative choices in selecting the appropriate CSR model to implement in their schools. More research and information about teaching practices and classroom interactions in CSR models is needed.

Wiley, Good, and McCaslin (2008) conducted an observational study in 16 Arizona schools using a systematic coding system to learn more about the teaching practices and classrooms interactions in schools implementing CSR models. They
focused on math and reading/language arts in classrooms implementing various models of CSR programs. Mathematics classes observed focused mainly on computations and math application activities. The reading/language arts classes focused on literacy related activities such as spelling, vocabulary, phonics, and writing development. The primary research questions guiding their study included (a) does subject matter matter in CSR classrooms, (b) how does instruction in CSR classrooms differ among Grades 3, 4, and 5, and (c) are there major differences in classroom practices between fall and spring?

The population of participants for the study included teachers ($N = 104$) in grades 3-5 in 16 CSR schools. This consisted of 248 observation periods of the 104 teachers in mathematics and reading/language arts classes. Analyses examined the effects of semester (fall = 169, spring = 79), grade (Grade 3 = 71, Grade 4 = 96, and Grade 5 = 81) and subject matter (math = 70 and reading = 178) on observed teacher practices. Wiley et al. (2008) used the Comprehensive School Reform Classroom Observation System (CSRCOS) to measure the presence or absence of instructional practices that characterize effective teaching. The observations included 10 minute observation intervals. The first five minutes focused on coding classroom activities. The next three minutes were for recording reflections and fine-tuning narratives. The last two minutes the observers focused on scoring the recorded intervals. Classroom practices were recorded in three categories including instructional opportunities, student activity, and teacher-student relationship.

The general trends presented in the data analysis supported previous studies in suggesting that CSR classroom practices are relatively stable across different school years. Students in this study (Wiley et al., 2008) appeared to be engaged in learning basic skills in an uninterrupted teacher directed classroom setting. Students did what was asked
of them by the teacher with little opportunity to make choices involving the social or academic tasks. The mathematics lessons observed appeared to be orderly and predictable while the reading lessons were more fluid. For this study the same teachers' instruction varied across different subjects which suggested strongly that differences are curricular or content related rather than teacher related.

In sum, Beam and Faddis (2001), Carlson (2003), Holdzkom (2002), Sterbinsky et al. (2005), and Ross et al. (2004) studied Comprehensive School Reform and found that implementation had a positive impact on school environment and student achievement. Schools that implemented each of the components of the school reform model experienced an increase in student achievement. Wiley, Good, and McCaslin (2008) and Zhang, Fashola, Shkolink, and Boyle (2006) examined the relationship of improvement over time and years of implementation on student achievement. They found that length of time in the program impacted student outcomes. These studies were important because they began looking at the impact of Comprehensive School Reform beyond the initial year of implementation. Since the grant funding for implementation of Comprehensive School Reform models lasted three years, no data were gathered beyond that time span.

The funding that schools receive to implement any Comprehensive School Reform model lasts for three years. During this time schools receive support to implement the components of the program effectively. To maintain the level of gains in student academic achievement beyond the initial three years, it is important for schools to continue implementation of the program relying only on internal support. Future research should address the long term impact of CSR after funding is no longer available.

Good et al. (2005) and Epstein (2005) studied the impact of Comprehensive
School Reform components and their impact on schools in various states. The research showed that focused implementation of the components of the Comprehensive School Reform model implemented in the school has a positive impact on student achievement. Comprehensive School Reform models have been researched and developed with the intent that specific components, when implemented together and effectively, guided schools toward the desired outcomes of increased student achievement. It is important to continue to implement all of the components together. When schools vary from this implementation the results they achieve are not consistent or sustainable.

Wetherill and Applefield (2005) and McCaslin et al. (2006) conducted research to determine the connection between student success and teacher implementation of comprehensive school reform models. Findings from their studies showed that schools achieving higher student success implemented the six basic components. Schools not implementing all components experienced lower levels of student success. It is important for administrators and teachers to become familiar with the components in order to implement them successfully and comfortably to achieve school wide success from the reform model.

Ross et al. (2004) and Carlson (2003) conducted research focused on the issue of sustaining the comprehensive school reform efforts after the initial implementation of various models. Since the models implemented impacted the schools successfully during initial implementation these researchers were interested in the level of success after the grant support ended and schools were left on their own to continue implementation. The results of their studies showed that schools that had stronger buy-in were able to sustain the effects and see noticeable gains on academic achievement after the initial implementation and support of the reform model ended. Both Ross et al. and Carlson
found that the determination of success of the programs depended on several factors: financial support, school support, and classroom change. These factors combined influence the interpretation of successful programs.

The data analyses from the study of these schools show that CSR orientation overall was associated with measurable changes in school climate, pedagogy, and student achievement. Several researchers (Ross et al., 2004; Carlson, 2003; Sterbinsky et al., 2005) all conducted research that links implementing various CSR models with a positive impact on student achievement. Effective implementation of all components of the CSR model selected had a positive impact on classroom practices and school climate which led to increased student academic achievement (Sterbinsky et al., 2003); however, the effect sizes varied considerably from one study to the next and some improvements were modest at best. Researchers attributed these variations to level and fidelity of implementation.

America’s Choice School Reform

Many administrators and educators believe that comprehensive school reform holds promise for improving low performing schools because these programs are research based, student-centered and have documented success. According to the literature, there are presently over 500 Comprehensive School Reform designs that are research based that simultaneously reshape many elements of education including curriculum, instruction, and school governance. One of these programs is the America’s Choice Comprehensive School Design. Federal grants known as the Comprehensive School Reform program spurred the growth of the America’s Choice comprehensive design which has been implemented in schools across the United States.

The America’s Choice (AC) Comprehensive School Design was developed by the
National Center on Education and Economy (NCEE), founded in 1989 as a non-profit organization, to address the notion that children in the United States must gain worldwide skills and knowledge to compete in the world economy (NCEE, 2002). In 2004, America’s Choice separated from NCEE and became a for profit institution. America’s Choice is the second generation of the NCEE National Alliance for Restructuring Education.

The America’s Choice school design is a comprehensive K-12 school reform model that is implemented in over 500 schools across our nation. America’s Choice focuses on raising academic achievement by providing rigorous standards-based curriculum and safety-nets for all students. The America’s Choice comprehensive design is based on research that influences leadership, teacher’s professional development, and curriculum. It relies heavily on student-performance standards which put the focus on students and their progress. The goal of America’s Choice is to make sure that all but the most severely handicapped students reach an internationally benchmarked standard of achievement in English language arts and mathematics by the time they graduate.

During implementation of America’s Choice, teaching staff receives training, coaching and support to implement Reader’s, Writer’s, Mathematics, and Science Workshops over a three year period. The first year involves America’s Choice staff who provide training and support for literacy coaches as they introduce Reader’s and Writer’s Workshop. The second year adds coaches for Math Workshop and introduces Science Workshop. In contrast, the other models of comprehensive school reform each had their own distinctive factors. Schools that chose America’s Choice decided that the options offered best met their needs.

In recent years many researchers have studied the various components, aspects,
and influences relating to implementation of the America’s Choice design in schools across America. The research is relatively new relating to this program and primarily looks at implementation of the elements and components and their impact on student performance and school climate. The elements and components are listed below:

1. **High Expectations** for student performance expressed in the *New Standards Performance Standards* that specify what students should know.

2. **Initial focus on literacy** during the 2 ½ hour literacy block including Reader’s and Writer’s Workshops to develop basic skills.

3. **Common core curriculum** that is aligned with the standards and incorporated within the workshop format.

4. **Standards-based assessments** provide detailed feedback to the teachers and students about student’s skill levels in relation to the standards.

5. **Distributed school leadership structure** led by the school’s principal to coordinate implementation, analyze assessment results, implement safety net program and aligns schedules and other supporting activities.

6. **Safety nets** are times structured into the daily schedule that provide extensive support for students who need further assistance to meet the standards.

7. **Commitment to teacher professionalism** by providing ongoing, on-site professional development training in which content, pedagogy, and the standards are connected.

8. **Workshop format** provides organized blocks of instructional time including whole group instruction, small group work, and one-on-one conferencing to help students.

9. **Professional learning communities** are established to provide teachers with an organized means to evaluate student work samples and discuss the connection to instruction. (Supovitz & May, 2003, p. 6).

Supovitz, Taylor, and May (2002) compared the writing, reading, and mathematic test scores of schools using the America’s Choice school design with schools that are not using this design. The research question asked if the America’s Choice school design impacted student achievement. For this study, the population included Duvall County,
Florida schools implementing and not implementing the America’s Choice school design. Participants were ten elementary schools and four middle schools in the Cohort II implementation phase of the design and 38 elementary schools and seven middle schools in the Cohort III implementation phase. Schools in the district not using the America’s Choice design were the control group.

Supovitz et al. (2002) computed descriptive statistics of America’s Choice schools against other schools in the district. Data collected consisted of student and school demographic information and student test scores in reading, writing, and mathematics. The researchers analyzed data from 23,000 fourth and fifth grade students from 101 elementary schools and 20,000 seventh and eighth grade students in 27 middle schools in Duvall County. Supovitz et al. excluded from the analysis 22-30% of the elementary students and 28-38% of the middle school students due to missing or incomplete information.

Supovitz et al. (2002) transformed the achievement scores into z scores for the data analysis in order to compare the magnitude of effects, regardless of test and subject area. Six student independent variables for the analysis included (a) prior standardized achievement score, (b) number of days absent, (c) dummy indicators for free or reduced lunch, (d) minority student, (e) male student, and (f) disability classification. Six school-level independent variables included (a) school size, (b) average class size, (c) school grade (A-F) assigned under state accountability system, (d) percentage of students qualifying for free or reduced lunch program, (e) percentage of students absent more than 21 days, and (f) dummy variable used for schools using the America’s Choice design. Findings showed few detectable differences between the standardized reading performance and mathematics scores but a difference in performance of elementary and
middle school students in the area of writing in schools implementing the America's Choice school design (Supovitz et al.).

During the initial years of implementation of a school reform model there is layered support consisting of training or professional development for teachers implementing the program. In America's Choice schools, the layered support consisted of Principal Academies and network meetings for school leaders, plus training for the literacy and mathematics coaches. This support was in place for three years. After the initial year's end, it is up to the teachers and staff to maintain the level of implementation without the levels of support. Teachers have to continue to implement all components of the school reform model to maintain the level of student growth. Where differences occur is when teachers do not consistently implement all components of the model implemented in their school. Supovitz and May (2004a) conducted a longitudinal study in Florida to explore the relationship between teachers' implementation of different components of America's Choice school design and the learning gains of their students. The study also examined whether teacher characteristics were associated with teachers' implementation of the different components of the America's Choice school design.

The target population for the two-year study included $N = 186$ general elementary, English/language arts, reading, or writing teachers and $N = 2,187$ first through sixth grade students in ten schools in Plainfield, New Jersey. The sample participants were $n = 1,572$ students and $n = 114$ teachers for whom there were valid data. Chi-square tests for independence for categorical variables and $t$ tests of mean differences for continuous variables addressed bias for missing data. These test collections adjusted $p$ values for multiple comparisons using the Bonferroni method. Data collection included student test scores from the New Jersey Goals Performance
Assessment (NJGOALS), and the Elementary School Proficiency Assessment (ESPA). All test scores were standardized to a mean of zero and a standard deviation of one in order to remove any artificial influences of differences in test score scaling (Supovitz & May, 2004a).

Data analysis revealed statistically significant advantage (one-sixteenth of a standard deviation gain in student learning) for upper elementary students in a class where the teacher reported higher implementation of America's choice components. Supovitz and May (2004a) reported that average gains in student learning occurred when teachers implemented at least eight components of America's choice Design.

Supovitz and May (2004b) systematically examined the relationship between teacher implementation of the America's Choice Comprehensive School Reform model and gains in student learning in an urban, at-risk school. The research question was does the degree of implementation impact the level of student achievement? Participants in the study were from the Plainfield, New Jersey school district, a high poverty, high minority district that has implemented the America's Choice school design. The district included ten elementary schools, two middle schools, and one high school. The 7,500 students were predominantly African American (71%) and Hispanic (28%). Of the student population, 65% qualify for the free or reduced lunch program. It was one of the poorest districts in the state of New Jersey.

Supovitz and May (2004b) collected data from the 114 general elementary language arts teachers and 1,572 first through sixth grade students from the ten America's Choice schools. Achievement test scores from the current and previous years were used to measure gains in reading achievement. Of the target group of participants, 1,398 were present in the district for both years. The teachers received surveys to complete for the
study. The researchers had a return rate of 89%. Teachers and students in a class had to be present for both years to participate in the study. Data were also collected from a larger sample of America’s Choice teachers from across America to compare the results.

Hierarchical linear modeling (HLM) was used to control the variables during the analysis of the data. The variables included gender, ethnicity, free or reduced lunch, and mobility for students. Gender, ethnicity, and years of experience were examined for teachers. Thus only the most stable families were included in the study. The investigation excluded those students from the most unstable families (more mobile, more intensive poverty). Accordingly, the study’s findings may not hold for schools or students with these more concentrated levels of poverty (see Wilson, 1987). Partition variance analysis separated the data into within school and between schools, examining the following variables: (a) impact of teacher’s overall implementation on student learning, (b) the relationship between implementation of writer’s workshop and student learning, (c) the relationship between implementation of reader’s workshop and student learning, (d) relationship between teacher beliefs and student learning, and (e) implementation variables related to student learning (Supovitz & May, 2004b). The findings support the claim that teacher implementation of a comprehensive school reform model is related to student gains. The data show that more variation occurs within schools than between schools. This supports the earlier research findings that more variance occurs between teachers than between schools (Beam & Faddis, 2001; Sterbinsky, Ross, & Redfield, 2003; VanMeter, 2005). However, a major limitation of this study is the bias inherent in the sampling frame because only those students who were present for both years of the study were included.

May and Supovitz (2006) conducted an 11 year longitudinal study of the impact
of America’s Choice school design on student learning gains in Rochester, New York. Their research was guided by three central questions based on the primary goal of the America’s Choice school design that all students meet or exceed the same high standards in reading and mathematics. The three research questions were (a) is there evidence that America’s Choice school design increases students’ rates of learning and, if so, how big is the increase, (b) does America’s Choice school design improve the performance of particularly low achieving students, and, if so, is this accomplished at the expense of the higher performing students, and (c) does America’s Choice school design make education more equitable for minority students and, if so, is this accomplished at the expense of nonminority students. Participants in the study included more than 55,000 students in grades one through eight in 42 elementary schools and ten middle schools who were tested in reading and mathematics.

During the 11 years of this study Rochester schools used different achievement tests including the Stanford Achievement Test (SAT-9), the California Achievement Test (CAT-%), the Degrees of Reading Power test (DPR), the New York State assessments (NYS), the New York Pupil Evaluation Program tests (PEP), and the New York Preliminary Competency Test (PCT). The SAT-9, CAT, and DRP are nationally normed standardized tests. The NYS, PEP, and PCT are New York State assessments. Since the majority of the test scores were from district wide administration of the SAT-9, all other test scores were vertically aligned to the same metric scale so that growth of student performance could be tracked over time. All of these assessments were similar in content, format, reliability, use, and consequences. Bloom’s work (cited in May & Supovitz, 2006) on interrupted time series was incorporated into Bayesian hierarchical growth curve analysis with crossed random effects to compare longitudinal gains in test
performance of students attending America’s Choice schools versus students attending other Rochester schools and students in the America’s Choice schools before implementation of the CSR model. The time series models used for their analyses were three-level hierarchical growth curve models.

Overall, the students who attended the America’s Choice schools in Rochester experienced significantly greater annual gains in both reading and mathematics performance when compared to students in other non America’s Choice schools in the district. Students in first through third grades experienced an additional $\frac{7}{10}$ of a month learning each year compared to similar students in other Rochester schools (e.g., $[2.0 ÷ 29.7] \times 10 = 0.7$). Students in fourth through eighth grades in America’s Choice schools showed an additional 1.7 months of learning each year in reading and 2.6 months of learning in mathematics when compared to students in schools not implementing the America’s Choice model. After concluding the data analyses for this study, May and Supovitz (2006) determined that on average, students in America’s Choice schools learned significantly more than did other students in the district. Each year the results were small to moderate but over time they accumulate. The impact is two to three times larger for students in later grades (1.7 months in reading and 2.6 months in mathematics in grades 4 through 8) than in the early grades (0.7 months in both reading and mathematics in grades 1 through 3). These findings are consistent with other research on America’s Choice.

The study provides strong evidence of longitudinal effects of America’s Choice school design. It does have two limitations. First, schools in Rochester were not randomly assigned to America’s Choice and second, the researchers did not attempt to connect improvement in performance directly to implementation of specific components of the
America's Choice school design. May and Supovitz (2006) hoped to show the importance of long term longitudinal research in education.

VanMeter (2005) conducted a quantitative exploratory study to investigate the relationship between America's Choice comprehensive school reform, years of implementation, and student achievement in Kentucky elementary schools. The three guiding questions for this study were (a) does a difference exist in reading achievement scores for students after one, two, or three years of implementation, (b) are the schools able to sustain the scores after three years of implementation, and (c) does a trend exist in reading scores after four years of America's Choice implementation? The participant pool for the study consisted of 1,850 third grade (exiting primary) students from nine elementary schools in Kentucky who implemented the America's Choice comprehensive school reform model. All of these students took the Comprehensive Test of Basic Skills (CTBS/5) annual achievement test for Kentucky. Five of the nine schools included in this study did not meet all of the requirements set by the researcher because they did not have the fourth year of data. Two of those five schools also lacked third year data. These five schools were still included to add information about trends that might help influence others considering the America's Choice school design.

The data analysis consisted of using t tests to examine the reading Comprehensive Test of Basic Skills (CTBS/5) scores for third grade exiting primary students in each of the nine schools beginning with the year before implementation of America's Choice school design and ending with the year after the third year of implementation. For the data analysis the independent variable was years of implementation and the dependent variable were the annually administered reading NCE scores on the Comprehensive Test of Basic Skills (CTBS/5). The results of the t test comparisons for the four schools that
met all the criteria for this study were divided. Two of the four schools continued to show a decline in reading achievement scores after four years of implementation, one school’s scores declined slightly between year three and year four of implementation, the final school continued to show improvement each year of implementation of the America’s Choice school reform model. The results from VanMeter’s (2005) study on the impact of America’s Choice school design on reading achievement indicate there is no conclusive evidence to support or deny a statistically significant impact. The study was limited in the fact that there are 1,271 schools representing 176 school districts in Kentucky and only four schools met the final criteria to be included in the data analysis for this study.

VanMeter’s (2005) study adds to the body of research about the America’s Choice comprehensive school reform model. Combining the results of this study with the body of other research on America’s Choice provides insight for administrators who are considering selecting this model of school reform for their school. However, the differential findings in VanMeter (2005) suggest that fidelity of implementation may be an important factor, i.e., doing the program well is different from putting the name of the program on school letterhead and proceeding to do nothing different. Real reform is hard work. Instructional practices, beliefs about poor children’s abilities to achieve well, aligning curriculum and instruction, holding both teachers and students accountable--these and the other components of comprehensive reform require significant commitment on the part of faculty and both internal and external support in order to change the existing learning climate and culture of the school, and to maintain these changes (institutionalize them) after the external support is withdrawn.

In sum, the America’s Choice school design attempts to restructure the instructional practices in the content areas over a three year time span. The America’s
Choice comprehensive reform focuses on establishing a framework based on national standards to enable all students to become successful in all areas of reading and mathematical concepts and skills. Effective implementation of all of the components impacts the academic achievement outcomes.

Supovitz et al. (2002) and Supovitz and May (2004a, 2004b) studies found that student gains were linked to levels of teacher implementation of the America’s Choice components. The positive impact begins slowly with the first year of implementation and gradually increases through the third year of implementation. Low performing and minority students benefit from the effective implementation of the America’s Choice design as evidenced by the May and Supovitz (2006) study that showed gains in academic achievement for these student populations. VanMeter (2005) looked more specifically at third grade reading scores in low performing schools. His results were inconclusive, with no significant findings. However, VanMeter noted the small number of schools with adequate data and suggested that degree of implementation of the America’s Choice program design is likely to make a difference in outcomes. In general, the research on comprehensive school reform suggests that better fidelity to the model increases the likelihood of having a significant impact on student achievement outcomes (May & Supovitz, 2006; Sterbinsky, Ross, & Redfield, 2005; Supovitz & May, 2004a, 2004b; Supovitz et al., 2002).

Mathematics Achievement

Every child has the capability to succeed in school yet far too many children fail to meet their potential, especially in the area of mathematics. As reform models were introduced to the educational community, many focused on literacy skills development and overlooked the area of mathematics skills development. The Agenda for Action:
Recommendations for School Mathematics (NCTM, 1980) caused educators to focus on mathematics instruction for all students. This was the beginning of a reform movement that included explicit mathematics curriculum and skills development.

An important consideration in the ensuing math reforms was equity. Over the years, achievement gaps in mathematics scores between various subgroups such as low socio-economic status (SES) and between minority groups grew larger. Lubienski (2000, 2001, 2002) noted that students of lower socio-economic status and minority groups have received more than their fair share of rote learning and low-level exercises from teachers who expect little of them. In order to address this disparity a change was needed in the approach to mathematics, to be focused on discussion, problem solving, reasoning, and thinking rather than computation based on rote rules.

Recommendations for reform in mathematics education uniformly call for an increased emphasis on meaningful experiences in mathematics and decreased emphasis on the repeated practice of computational algorithms (Hiebert, 1992; Kenny & Silver, 1998; NCTM, 1989; Palacios, 2005; Wearne & Hiebert, 1998). This change in focus of instruction called for a drastic restructuring of traditional mathematics curricula. Textbooks were designed with the assumption that new teaching materials facilitated the shift from an algorithmic approach to teaching mathematics to a more conceptual approach (Ball & Cohen, 1996; Cobb, 1999; Hiebert & Wearne, 1993; Hiebert et al., 1998; Manouchehri, 1998; Manouchehri & Goodman 1998; Reys et al., 1997; Van Haneghan, Pruet, & Bamberger, 2004).

Due to the increased level of mathematical literacy necessary for everyday life in our increasingly global economy, many students suffer because of their lack of mathematical fluency and abilities (Kenny & Silver, 1998). Studies indicate (e.g.,
Bodovski & Farkas, 2007; Corwin & Storeygard, 1995; Elliott, 1996; Hiebert & Wearne, 1993) that a strong foundation in mathematics—cultivated in the early grades, when children need to develop basic math aptitudes and the critical thinking skills needed to succeed in algebra—is absolutely essential if students are to succeed in college and the workplace.

**Mathematics Standards**

The mathematics education reform movement has been underscored in several documents over the last 20 years including: *An Agenda for Action: Recommendations for School Mathematics* (NCTM, 1980), *Curriculum and Evaluation Standards for School Mathematics* (NCTM, 1989), *Professional Standards for Teaching Mathematics* (NCTM, 1991), *Assessment Standards for School Mathematics* (NCTM, 1995), and *Standards for School Mathematics* (NCTM, 2000b). The vision of the standards movement calls for mathematics teachers to rethink the nature of school mathematics, moving away from the subject as absolute content with definitions towards math as a discipline with greater emphasis on conceptual learning, reasoning, and the importance of relevant connections.

Recognizing that the majority of students nationwide were not learning mathematics with depth and understanding and that teachers were not engaging students in mathematical thinking and problem solving, the National Council of Teachers of Mathematics (NCTM, 2000a) set forth the *Principles and Standards for School Mathematics* (PSSM) to serve as a foundation for the improvement of mathematics curricula, teaching, and assessment. PSSM is divided into six principles that describe specific components and characteristics crucial to the development of a strong mathematics program:
1. Equity. Excellence in mathematics education requires equity--high expectations and strong support for all students.

2. Curriculum. A curriculum is more than a collection of activities: it must be coherent, focused on important mathematics, and well articulated across the grades.

3. Teaching. Effective mathematics teaching requires understanding what students know and need to learn and then challenging and supporting them to learn it well.

4. Learning. Students must learn mathematics with understanding, actively building new knowledge from experience and prior knowledge.

5. Assessment. Assessment should support the learning of important mathematics and furnish useful information to both teachers and students.

6. Technology. Technology is essential in teaching and learning mathematics; it influences the mathematics that are taught and enhances student's learning. (NCTM, 2000a, pp. 3-4)

Even though PSSM (NCTM, 2000) provided a framework for a new direction in mathematics education, Hiebert (2003) and Van De Walle (2006) point out that the traditional pedagogy of training students in computational procedures of arithmetic and algebra continues to prevail in many classrooms across America. The mathematics component of the America's Choice comprehensive reform model addresses these concerns by providing opportunities for active student engagement during mathematics workshop. Students are to develop thinking and problem solving skills as they learn the content and skills associated with mathematics.

Mathematics Curriculum

There is an increasing concern for the lagging mathematics performance of U.S. students compared to other nations. The Trends in International Mathematics and Science Study (TIMSS) provides data on mathematics and science achievement for fourth and eighth grade students in the United States compared to other countries as tested in 1995,
1999, 2003, and 2007. The 2007 report focuses on the performance of U.S. students in relation to their peers from other countries and on changes in mathematics achievement since 1995. The report also provides additional information about trends in the United States regarding achievement of students by sex, race/ethnicity, and enrollment in schools with different levels of poverty. According to the TIMSS report, the average mathematics score of U.S. fourth grade students was higher than those in 23 of the 35 other countries, lower than eight countries (mostly Asia or Europe), and no measurable difference from the remaining 4 countries that participated in the testing (Gonzales et al., 2008).

To ensure our nation's future competitiveness, economic viability, and security, President Bush created the National Mathematics Advisory Panel (NMAP) in April 2006. This panel examined the best scientific evidence on improving the teaching and learning of pre-K through 8th grade mathematics and subsequently presented 45 findings and recommendations to ensure all American students are prepared for and successful in learning algebra. The panel’s work updated earlier work (NCTM, 1991). In general these recommendations focused on curricular content, learning processes, teachers and teacher education, instructional practices, and instructional materials (NMAP, 2006).

In addition to mathematical understanding and skills development, students need to develop their abilities to analyze problems and to communicate ideas related to problem solving. The focus in mathematics curriculum reform has been a shift towards supporting the development of reasoning and thinking skills in addition to specific mathematics content (Elliott, 1996; Hiebert et al., 1998; McCrone, 2005). For example, mathematical manipulatives, often referred to as hands-on learning, can play an instrumental role in assisting students in developing a deeper understanding of concepts and skills (Sowell, 1989). Researchers have documented that instruction that centers on
the use of manipulatives produces greater mathematical gains and achievement as compared with instruction not using them (Fuson, 1992; Wearne & Hiebert, 1998).

The National Science Foundation (NSF) is an independent federal agency funded by Congress in 1950 to promote progress and advance the nation in the fields of science, mathematics, computer science, and other social sciences. The NSF offers the Math and Science Partnership (MSP) program as a response to the national concern for increasing educational performance of students in the areas of mathematics by providing funding support for K-12 math curriculum projects in an effort to produce programs that reflect the vision of the standards. At the elementary level, these projects include Everyday Mathematics, Math Trailblazers, and Investigations in Number, Data, and Space (Reys, 2001). Many of these are utilized as part of reform initiatives.

One problem with curriculum reforms is that it often occurs at the macro level (professional associations, state guidelines) and does not trickle down to the classroom level. In general the developers of comprehensive reform models recognize this and incorporate specific curriculum guidelines into their models. That was in the case with America’s Choice, which utilizes a math workshop framework for the math component to provide differentiated instruction to enhance learning. With a focus on mastering concepts, skills and problem solving, students incorporate the use of manipulatives and discussions to enhance their learning and understanding of mathematics during math workshop.

Classroom Practices

According to Palacios (2005), most students can demonstrate competency with basic mathematical facts but experience difficulty with their application to problem-solving situations. Analyzing data from the National Assessment of Educational Progress (NAEP), Berry (2003) identified teacher instructional methodology and practices as one key factor
that contributed to the poor achievement levels in students’ mathematics literacy.

One approach to instruction in mathematics involves students learning together. With the recent emphasis on providing opportunities for students to communicate about mathematics, cooperative learning strategies have taken on new significance (Corwin & Storeygard, 1995; Gadeyne, Ghesquire & Onghena, 2006; Lubienski, 2000; Means & Knapp, 1991; Riordan & Noyce, 2001). The interactions that occur in cooperative learning are consistent with the small-group setting which appears to provide a natural environment for increased dialogue and communication. The discussions about mathematics which occur among students can help them internalize conceptual understanding; however, Manouchehri (2004) suggests that the small-group activities must be structured to maximize the chance that students engage in questioning, elaboration, explanation, and other verbalizations in which they can express their mathematical problem solving and thinking, precisely the focus in cooperative learning. The students must move beyond just answering problems.

Projects such as Summer Math for Teachers (Schifter & Fosnot, 1993; Schifter & Simon, 1992; Simon & Schifter, 1991), the Purdue Problem Centered Mathematics Project (Cobb et al., 1991; Wood, Cobb, & Yackel, 1991), and Cognitively Guided Instruction (Carpenter & Fennema, 1992; Fennema et al., in press), focused on teachers’ concepts of mathematical learning as a basis for helping educators make fundamental changes in instruction. The assumption of all three projects is that students construct knowledge rather than simply assimilate some parts of what they are taught (Cobb, 1994; Davis, Maher, & Noddings, 1990). It is believed that children enter school with informal knowledge about mathematics that can provide the basis for developing their understanding about basic math skills and concepts. Through Cognitively Guided
Instruction (CGI), students move from applying a formula to solve a problem to finding their own solution and justifying or explaining the solution to their friends and the teacher.

While math curriculum and textbooks show changes directly connected to the math reform movement, minimal changes in how math is taught are reflected in classroom practices. A number of studies (e.g., Carpenter & Moser, 1984; Carpenter, Hiebert, & Moser, 1981; Riordan & Noyce, 2001) have demonstrated that semantic structure is relevant in studying children's solutions of word problems. Carpenter and Moser (1983) developed a classification of word problems based on semantic structure: Change, Combine, Compare, and Equalize. The skills developed progress from the easiest (Change and Combine) and progress to the most difficult (Compare and Equalize). Research has shown that skill in solving word or story problems gradually increases during elementary school (Riley & Greeno, 1988). The NCTM Standards recommend that curriculum and classroom practices should place an emphasis on problem solving, reasoning, making connections between mathematical topics, communicating mathematical ideas, and providing opportunity for all students to learn (NCTM, 1989, 1991, 1995, 2000a). Several empirical studies have been conducted on the effects of those practices.

McCrone (2005) conducted a qualitative observational study using mixed methodology to examine how mathematical discussions develop and to what extent the teacher's role impacted the development. The participants selected for the study were from one fifth grade classroom in a school that borders a large metropolitan city in the northeastern United States. McCrone observed the mathematics class over a six month period. Two questions guided the research; (a) How do student contributions to
discussions develop over time, and (b) What pedagogical choices did the teacher make that related to the mathematics discussions? The teachers’ focus in this classroom was to establish mathematical inquiry through discussions as the basis for student learning.

Data collection consisted of videotapes, audiotapes, field notes, interviews, and a shared reflective journal. Data collection began the first six weeks of school with weekly follow-ups one week a month for the remaining months of the school year. A modified version of Miles and Huberman’s three part analysis (cited in McCrone, 2005) was used to organize and analyze the data. This three part analysis consisted of (a) developing a comprehensive description of the observations, (b) classifying the data recorded into categories to assist with providing a response to the research questions, and (c) evaluating the data to look for ways that the data are interconnected and impact the study. The researcher focused the data analysis on two categories: patterns of discourse and levels of response.

The study showed that patterns of discourse increased by the end of the year. At the beginning of the study whole group discussions were not as rich as those in small groups. By the end of the study whole class activities included rich conversations of mathematics. Findings also showed that the levels of response shifted from students verbalizing to justify solutions at the beginning of the year to verbal responses that included descriptions of solution processes and the reasoning involved in the steps by the end of the year. The nature of student contributions paralleled changes in the nature of student interactions. McCrone (2005) noted that by the end of the year the frequency of structure-oriented questions increased. The researcher concluded that the nature of the teacher’s questions and the expectations of student responses influence mathematical discussions. This investigation supports the idea that students are dependent on teachers
as models for developing the appropriate forms of communication in classroom settings.

Bodovski and Farkas (2007) used the Early Childhood Longitudinal Study – Kindergarten Cohort (ECLS-K) to examine the gains in mathematics achievement in the first four years of elementary school. The ECLS-K, sponsored by the U.S. Department of Education National Center for Statistics, selected a nationally representative sample of kindergarten students \((N = 17,487)\) in the fall of 1998 and followed them through the end of eighth grade. The participants for this study \((n = 13,043)\) consisted of students who had mathematics data available at four times: fall of kindergarten, spring of kindergarten, spring of first grade, and spring of third grade. This study had the following three guiding questions. First, does mathematics knowledge at the beginning of kindergarten affect a student's rate of growth in the following years? Second, what is the effect of time spent on math instruction on student achievement? Third, what is the effect of student engagement with learning on achievement growth?

Bodovski and Farkas (2007) conducted a two part data analysis. The first part analyzed three models for each achievement gain and the second part looked in great detail at student’s mathematics skills in the fall of kindergarten. The students were placed into four groups based on the results of their scores on the initial mathematics assessment given in the fall of their kindergarten year. The mathematics scores for this study were measured using item response theory (IRT) so the results were comparable over time. The data analysis showed that the student’s beginning knowledge was associated with how much they gained during the first four years of elementary school. The lowest group (below the 25th percentile) gained 56.3 points by the spring of third grade. The second group (25th to 50th percentile) gained 62.5 points. The third group (51st to 75th percentile) gained 67.2 points and the fourth group (above 75th percentile) gained 66.8
points. The students in the lowest group received the greatest amount of mathematics instruction. The higher achieving group had higher student engagement. In sum, Bodovski and Farkas (2007) found that student’s mathematical proficiency when they entered kindergarten is consequential for later mathematics achievement growth. Also, the low school engagement of students in the bottom quartile plays an important role in the lower achievement levels for this group of students. This research mirrors the findings of Berry (2003) on NAEP data, that instructional practices contribute to poor achievement particularly those with the lowest scores.

In summarizing these studies, McCrone (2005) and Bodovski and Farkas (2007) studied various aspects of effective practices in mathematics instruction and their impact on student achievement gains. Both investigations reported on the impact of the teacher’s role, time spent on daily mathematics instruction, and the impact of student engagement. Providing opportunities for students to engage in discussions about mathematical concepts through structured cooperative learning strategies helps students internalize understanding of math concepts. To be effective, these small group interactions should engage students in questioning, explaining, and talking about the problem solving skills used in math (Manouchehri, 2004). All of these are important components of a sound mathematics program. When considering adopting a school wide reform model, it is important to evaluate the components of the mathematics instructional program and determine what components should be the focus of change. This information helps in the selection of an effective school reform model. Another factor to consider is how the system selected will be monitored for student achievement progress.

Achievement Gaps

Current reforms in mathematics are intended to decrease past inequities or gaps in student
achievement and instructional opportunities by offering mathematics education focusing on developing the skills of problem solving and critical thinking (Corwin & Storeygard, 1995; Englert, Barley, Apthorp, Lauer, & Van Buhler, 2005; Lee, 2006; Lubienski, 2001; Lubienski & Shelley, 2003). In order to increase student achievement in mathematics, schools must look closely at several factors relating to instructional practices in their building. Closing achievement gaps is one such factor and the role of teachers in the mathematics classroom is another. The teachers play an important part in the daily instructional practices in the mathematics classrooms in each school. Several studies demonstrate that students who begin with the least mathematics knowledge also showed the least growth (e.g., Bodovski & Farkas, 2007).

Passage of the No Child Left Behind (2002) legislation introduced challenges to the mathematical community, causing educators to revisit their commitment to attaining equity in achievement outcomes for all students. While math initiatives introduced over the past 20 years included a focus on removing achievement disparities, this message is not apparent in classrooms with minority students, multi-lingual students, and students with disabilities (Matthews, 2005). In order to meet the guidelines set forth by No Child Left Behind, math reform efforts focus on the concept of a math-for-all approach to address achievement disparities among populations. Several researchers have investigated the issue of achievement gaps in curricular areas including mathematics.

One such longitudinal study by Lubienski and Shelly (2003) expanded on previous research related to race-related achievement gaps focusing on trends related to socioeconomic status (SES). The researchers investigated patterns in student mathematics performance and instruction, proposing that racial achievement gaps could be related to differences in student’s access to empowering mathematical instruction. Four questions
guided the study: (a) How have mathematics achievement gaps involving white, black, and Hispanic students changed over the past decade, (b) Are current race-related achievement gaps consistent across SES and gender groups, (c) Are reform-based instructional practices reaching all students regardless of race, and (d) For instructional practices implemented more with white students than black or Hispanic, to what extent do student and school SES account for the instructional disparities?

Lubienski and Shelly (2003) accessed mathematics assessments and survey data from the 1990, 1996, and 2000 National Assessment of Educational Progress (NAEP) using their CD-ROM data tool. Sample data for the study consisted of 8,072 students equally divided between 4th, 8th and 12th grades in 1990. In 1992 and 1996 the samples each had 21,000 students. The 2000 sample size doubled to 42,000 students. The researchers examined teacher and student questionnaire data to examine the instructional practices occurring in the mathematics classrooms. The data analysis consisted of cross tabulation to calculate means and standard errors for student achievement data and student and teacher questionnaire data about instructional practices. The analyses focused on descriptive statistics related to race-related differences that occur across school and student level SES groups. The researchers also utilized Hierarchical Linear Modeling (HLM) to determine the strength of the correlation between various demographic variables, student achievement, and instructional practices.

The results from the data analysis showed that mathematics achievement scores did increase between 1990 and 2000 for white, Hispanic, and black students and for low and high SES students' even though substantial gaps remain and are growing (Lubienski & Shelly, 2003). Several instructional differences were identified that could relate to achievement disparities. The findings show that white, high SES students are
experiencing more of the recommended instructional practices. Using the NAEP 500 point scale, the 2000 results showed 4th graders on average scored 228, 8th graders scored 275, and 12th graders scored 301. The 2000 Hispanic-white gap was 24 points at fourth grade, 33 points at 8th grade, and 25 points at 12th grade. The black-white gap was 31 points at fourth grade, 39 points at 8th grade, and 34 points at 12th grade. The Hispanic-white gaps are large but not as large as the black-white gaps.

Mathematics Reform

Ysseldyke and Tardrew (2007) conducted a study to investigate the effects of using a progress monitoring system (Accelerated Math) in elementary, middle, and secondary school settings. Accelerated Math is an integrated technology-based monitoring system that is intended to accelerate the mathematics learning when used to support existing instructional practices and mathematics curriculum in the classrooms. STAR Math is a computer adaptive test of skill development in mathematics.

Ysseldyke and Tardrew (2007) conducted the study in 125 classrooms (67 experimental and 58 comparisons) in 47 schools in 24 states during the spring semester of the year, January through May. A total of 2,397 students (1,319 experimental and 1,078 comparison) in grades 3 through 10 participated in the study. After considering several requirements the final sample consisted of 2,202. All schools were similar in major demographic variables. All of the students in both the experimental and control groups were pretested using STAR math, a computer-adaptive mathematics assessment in January and post-tested in May of the same year. For the data analysis, the researchers focused primarily on grades three through six. Researchers conducted an ANCOVA comparing post-test NCE scores earned on STAR Math by students in the experimental and control groups. Students in the Accelerated Math program in grades three through six
significantly outperformed the students who did not participate in the program ($p < .001$ at Grades 3-5, and $p < .05$ at Grade 6). A teacher survey was administered at the conclusion of the study to gather information about the instructional practices used by the teachers, homework assigned, and their views on whether students were profiting from mathematics instruction. Sixty-one teachers who used the progress monitoring system and 47 control group teachers returned the survey.

The results of the data analysis found the semester long implementation of a progress monitoring and intensive management system had significant positive effects on the performance of students in grades three through six. Students who participated in Accelerated Math out performed students in the control group (Ysseldyke & Tardrew, 2007). This study is limited by the fact that the program was put in place mid-year and therefore used only one semester, demonstrating the impact of implementing a focused system of progress monitoring student on mathematics achievement. This is one component of a mathematics program that enabled schools to select and implement a school-wide program that produced the desired effects. Another component of a successfully implemented program is based on sound mathematics curriculum implementation in the school.

MacIver and MacIver (2008) conducted a study in 86 Pennsylvania schools examining the relationship between mathematics achievement growth for middle-grade students on the Pennsylvania system of School Assessments and the number of years schools implemented either a Whole School Reform (WSR) model with National Science Foundation (NSF) supported mathematics curriculum or a WSR model without a mathematics curriculum component from 1997 to 2000. The focus of the study, which lasted four years, was to determine if students at schools with a WSR strategy
emphasizing the implementation of a coherent mathematics curriculum would show
greater mathematics achievement gains than students at schools with no specific WSR
strategy.

MacIver and MacIver (2008) used the hierarchical linear model (HLM) estimates
for the spring 2000 cohort of eighth grade students. Students’ scores from grade five were
used as a baseline to determine mathematics achievement gains. The mathematics
achievement gains for students were positively related to the number of years those
schools implemented a specific mathematics curriculum reform. Students at these schools
out-gained students at non-implementing schools by 1.12 NCE points for each year of
implementation of the reform model. The results of the study show that sustained
implementation of math-oriented whole school reform models is related to higher growth
in middle grade mathematics achievement among students in high poverty urban districts.
The overall effect sizes ranged from roughly 0.2-0.3 SD. The positive relation to student
achievement growth is found when schools implemented a WSR strategy using
standards-based mathematics curriculum and when they implement this model for
multiple years. The study is limited in the findings by the fact that the researchers were
unable to distinguish between different types of mathematics models and curricula
implemented in the schools during the study.

Ysseldyke and Tardrew (2007) and MacIver and MacIver (2008) studied whole
school reform models in relation to mathematics instruction. Both studies concluded that
the length of time schools implemented a school-wide reform model directly impacted
the student mathematics achievement results. Ysseldyke and Tardrew reported that using
computer adaptive mathematics programs--STAR and Accelerated Math--that included
progress monitoring and intense management to supplement the mathematics curriculum
had a positive impact on student achievement when compared to students in the control groups. MacIver and MacIver found that mathematics achievement gains were positively related to the number of years the school implemented the whole school reform model.

Recommendations for reform in mathematics education uniformly call for an increase in the development of conceptual experiences and a decrease in repeated practice in computational algorithms (National Council of Teachers of Mathematics, 1989, 1995, 2000). This trend is leading low performing schools to examine more carefully various models of school wide reform in mathematics. District administrators must consider several factors before selecting the appropriate model that will match the school culture to lead to effective implementation and the desired gains in mathematics achievement for their students and school. Student achievement in mathematics increases when schools change the focus of their instructional practices from computational skills to developing critical thinking skills in mathematics (Lubienski & Shelly, 2003; McCrone, 2005).

Current Trends and America’s Choice

In the current study the researcher looked at the impact of implementing the America’s Choice comprehensive school reform design for seven Kentucky elementary schools in the area of mathematics achievement for third grade exiting primary students. The study examined mathematics achievement scores for the year prior to implementation, two years during supported implementation of the America’s Choice comprehensive reform design, and the year after support ended to determine the impact of this program on each of the seven schools.

Literature that relates to this study includes research in several areas. In the area of school reform, work conducted by McCombs (2002), Dimmock (2002), Michael and Young (2005), and Peterson et al. (1996) is relevant. McCombs (2002) and Dimmock
conducted studies that developed and tested methods to assess the effectiveness of school reform models. Their studies showed that implementation and evaluation of research-based, student-centered school reform models assisted schools in achievement of the desired outcomes. Peterson et al. (1996) and Michael and Young (2005) focused their work on school restructuring from the administrative and teacher's point of view. Their research demonstrated that taking time to implement school reform correctly produces positive outcomes.

Studies by Beam and Faddis (2001), Holdzkom (2002), Good et al. (2005), and Epstein (2005) all addressed the impact of comprehensive school reform. Beam and Faddis (2001) reported findings from the Comprehensive School Reform (CSR) Implementation survey which revealed a high level of agreement among principals and teachers in regards to the degree of focus on development of all components and progress towards accomplishing their desired goals. Studies by Good et al. (2005), and Epstein (2005) focused their research on implementation of the components and found that intentional implementation guided schools to increased student achievement. Holdzkom (2002) studied change in student outcomes, teachers' practice, and teachers' perceptions that characterize the schools. The findings demonstrated that evaluating the reform model should include examining the components as well as the strategies addressed.

In recent years, Lubienski (2000, 2001, 2002), Corwin and Storeygard (1995), Manouchehri (2004), and MacIver and MacIver (2008) have conducted studies on various aspects of mathematics instruction and the impact on academic achievement. A change in the focus of mathematics instruction from rote learning to more focus on discussions, problem solving, reasoning and thinking was needed to address the disparity for low socio-economic and minority groups (Lubienski, 2000, 2001, 2002).
Manouchehri (2004) conducted research that supported the inclusion of structured, small group activities to engage students in questioning, elaboration, and explanation of mathematical thinking and problem solving. MacIver and MacIver (2008) studied whole school reform focusing on mathematics instruction and found that mathematics achievement gains were positively linked to the number of years of implementation of the comprehensive reform model. Corwin and Storeygard (1995) conducted research on the impact of increasing mathematical fluency and abilities to decrease inequities in student achievement which led to new significance in structured cooperative learning strategies. All of these works support the math reform movement’s focus on providing opportunities for students to engage in mathematical discussions to help conceptualize the main concepts of math.

Finally, several researchers, including Supovitz et al. (2002), Supovitz and May (2004a, 2004b), and May and Supovitz (2006), conducted studies examining the impact of America’s Choice comprehensive reform model in schools in various states. Findings from the Supovitz et al. (2002) and Supovitz and May (2004a, 2004b) studies showed that teacher implementation of America’s Choice components were linked to positive gains in student achievement beginning in the first year and continuing through the third year of implementation. Minority and low performing students also benefit academically from effective implementation of all components of America’s Choice reform (May & Supovitz, 2006).

The current study is different from the previous studies because the researcher focused on third grade exiting primary mathematics achievement for seven Kentucky schools implementing America’s Choice comprehensive reform design. Other studies have looked at the impact of America’s Choice in reading and writing or just reading.
The work by VanMeter (2005) comes closest to the current analysis. VanMeter examined nine elementary schools in Kentucky that implemented America’s Choice, with reading as the criterion. The results were inconclusive showing reading improvement evident in some schools but declining in others during each year of implementation. In contrast, this study revisits that data base to investigate impact on third grade mathematics scores.

Summary

For several decades, the federal government and American educators felt an urgency to improve academic achievement and student performance in American schools. Legislators and policy makers addressed these concerns by calling for comprehensive school reform. The report, *A Nation at Risk* (National Commission on Excellence in Education, 1983), focused on preparing students to become effective members of a global society by focusing on raising academic achievement.

Initiatives such as the Elementary and Secondary Education Act (ESEA) from 1965 and the more recent version of the No Child Left Behind Act of 2001 provided guidance for school districts and educators to implement changes leading to stronger literacy, science, and mathematics curriculum. These initiatives included funding to support professional development, instructional materials and resources to support educational programs, plus an emphasis on parent involvement. Many school reform models focused on one specific area but schools that were in need of school wide improvement needed to implement a more comprehensive model of reform.

Comprehensive school reform holds promise for improving student achievement because it aligns all parts of the system including standards, curriculum, instruction, assessment, classroom management, professional development, parent involvement, and school management (Ross & Gill, 2004; Rowan et al., 2004; Schmoker, 1996).
Implementing this systematic change can lead to success in student achievement. In order to affect the overall school environment, research-based programs that include instructional practices in all areas must be incorporated to achieve the desired outcomes of increased student achievement.

America’s Choice is one research-based comprehensive reform model that was developed by the National Center on Education and the Economy (NCEE) through the use of grant funds available from the Comprehensive School Reform program. America’s Choice is a comprehensive reform that focuses on raising academic achievement by providing rigorous standards-based curriculum and safety nets for all students (Supovitz et al., 2002). As part of America’s choice, teaching staff receive training and coaching support to implement Reader’s, Writer’s, Mathematics, and Science Workshops over a three year period.

Many of the reform models adopted by the educational community center on the development of literacy skills and overlooked mathematics skills. The Agenda for Action: Recommendations for School Mathematics (NCTM, 1980) led to a nationwide focus on mathematics instruction. The movement towards math reform calls for moving away from repeated practice of computational algorithms and providing more opportunities to develop reasoning and problem solving, while stressing the importance of making relevant connections (Hiebert, 1992; Kenny & Silver, 1998; NCTM, 1989; Palacios, 2005; Wearne & Hiebert, 1998). Implementing reform models that include a focus on mathematics addresses the concern of lagging mathematics performance of U.S. students when compared to other nations.

Several documents provided guidance and led the mathematics reform movement. Paving the way for a change in mathematics instruction in American schools were
resources such as: An Agenda for Action: Recommendations for School Mathematics (NCTM, 1980), Curriculum and Evaluation Standards for School Mathematics (NCTM, 1989), Professional Standards for Teaching Mathematics (NCTM, 1991), Assessment Standards for School Mathematics (NCTM, 1995), and Standards for School Mathematics (NCTM, 2000b). This shift in focus towards thinking, reasoning, and internalizing (constructing) mathematical insights, especially in the early years, will better prepare U.S. students to become meaningful and productive members of a global society.
CHAPTER III

METHODOLOGY

Introduction

Research on comprehensive reform suggests that the community, parents, students, administrators, faculty, and staff at consistently low-performing schools can implement school-wide efforts successfully if they have made a research-based selection and have support during the years of implementation. There needs to be collaborative team effort that inquires into the best practices for accomplishing their collective aims and outcomes for the organization or school. Any discrepancy between best practices and reality of their school culture or environment can lead to a reduction in the effectiveness of the program implementation (Holdzkom, 2002).

The purpose of this study was to examine the impact of implementing the mathematics component of the America’s Choice comprehensive school reform model over time for third grade exiting primary students in Kentucky. The investigation builds upon the work of VanMeter (2005) who studied the impact of America’s Choice on reading in the same set of schools.

In Kentucky, the America’s Choice comprehensive school reform model was implemented for yearly cohorts from 2001 to 2004. Nine Kentucky elementary schools opted for America’s Choice, one of several whole school comprehensive reform packages that schools could select. Each school applied for and received a Title I grant to implement the America’s Choice model available to low achieving schools through the Kentucky
Department of Education (KDE). As part of the implementation of this grant, the schools received support and professional development for three years. Building personnel had the option of continuing with the implementation for subsequent years without the additional funding.

This study addresses only the part of the program designed to improve mathematics, specifically, the Normal Curve Equivalent (NCE) mathematics scores from the Comprehensive Test of Basic Skills (CTBS/5) administered to third grade exiting primary students. The remainder of this chapter is divided into sections that provide information relating to the methodology of the study. Section one describes the Study Design, followed by information about the Population and Sample, including demographics and size of the schools. Sections on Data Analysis techniques and Description of the Variables are then explicated. Finally, two last sections address Validity and Reliability and then Ethical Considerations. The chapter concludes with a Summary.

Study Design

The study involves a longitudinal research design with a cohort model. A cohort study constitutes a comparative and observational design in which subjects are grouped by their exposure status, i.e., whether or not the subject was exposed to a particular level of the study factor. The focus of this research is the set of schools that implemented America’s Choice comprehensive school reform in Kentucky for the years beginning 2000-2001 through 2004-2005 (henceforth 2001-2005). Table 1 lists the nine Kentucky America’s Choice schools and the years that each implemented the program. The chart also lists the years prior to (beginning in 2000) and after the implementation years. Not all of the schools began implementation in the same year.
Table 1

*Years of Implementation of the America’s Choice Model in Kentucky Schools*

<table>
<thead>
<tr>
<th>School</th>
<th>P1</th>
<th>Y1</th>
<th>Y2</th>
<th>Y3</th>
<th>C1</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>B</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>C</td>
<td>2000</td>
<td>2001</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
</tr>
<tr>
<td>D</td>
<td>2001</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
<td>2005*</td>
</tr>
<tr>
<td>E</td>
<td>2001</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
<td>2005*</td>
</tr>
<tr>
<td>F</td>
<td>2001</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
<td>2005*</td>
</tr>
<tr>
<td>G</td>
<td>2000</td>
<td>2001</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
</tr>
<tr>
<td>H</td>
<td>2000</td>
<td>2001</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
</tr>
<tr>
<td>I</td>
<td>2000</td>
<td>2001</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
</tr>
</tbody>
</table>

*Note.* P1 = year prior to implementation; Y1 – Y3 = year 1 – 3 of implementation, respectively; C = continuation of CSR in year after active implementation ended; N/A refers to schools that did not implement America’s Choice for various reasons.

*This school did not actively implement America’s Choice during this year.*

For this study, the researcher used quantitative test data collected from the Comprehensive Test of Basic Skills (CTBS/5) in the form of Normal Curve Equivalency (NCE) scores from each of the schools that implemented the America’s Choice comprehensive school design. The data represented student level scores for the successive third grade cohorts of the years in question. Even though several schools did not participate in America’s Choice for their continuation year, the researcher obtained data from the state for those schools for that year. The information for each school was collected over a four year period including the year prior to implementation, the two years of implementation,
and the following year of continuation. Because of the staggered start to America’s Choice, this represented five years of data overall. Kentucky’s Commonwealth Accountability Testing System (CATS) occurred during the spring of each year for all schools. The use of the cohort data structure indicates that a different sample was collected at each subsequent data collection point. Thus, this study does not follow the same individuals over time. However, the population remains the same (i.e., all third grade students in Kentucky schools with CTBS/5 mathematics NCE scores for the years 2001 through 2005).

Table 2

Descriptive Characteristics for America’s Choice Schools in Kentucky for 2000

<table>
<thead>
<tr>
<th>School</th>
<th>Grades</th>
<th>Pop</th>
<th>%F/RL</th>
<th>%W</th>
<th>%ELL</th>
<th>Title I</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>K-5</td>
<td>231</td>
<td>98</td>
<td>39</td>
<td>1</td>
<td>X</td>
<td>rural</td>
</tr>
<tr>
<td>B</td>
<td>PreK-5</td>
<td>270</td>
<td>95</td>
<td>50</td>
<td>19</td>
<td>X</td>
<td>urban</td>
</tr>
<tr>
<td>C</td>
<td>PreK-5</td>
<td>600</td>
<td>95</td>
<td>90</td>
<td>10</td>
<td>X</td>
<td>rural</td>
</tr>
<tr>
<td>D</td>
<td>PreK-5</td>
<td>312</td>
<td>93</td>
<td>96</td>
<td>3</td>
<td></td>
<td>rural</td>
</tr>
<tr>
<td>E</td>
<td>K-5</td>
<td>312</td>
<td>66</td>
<td>95</td>
<td>4</td>
<td></td>
<td>rural</td>
</tr>
<tr>
<td>F</td>
<td>K-5</td>
<td>642</td>
<td>38</td>
<td>96</td>
<td>3</td>
<td></td>
<td>rural</td>
</tr>
<tr>
<td>G</td>
<td>K-5</td>
<td>335</td>
<td>95</td>
<td>99</td>
<td>1</td>
<td>X</td>
<td>rural</td>
</tr>
<tr>
<td>H</td>
<td>K-5</td>
<td>213</td>
<td>95</td>
<td>99</td>
<td>0</td>
<td>X</td>
<td>rural</td>
</tr>
<tr>
<td>I</td>
<td>K-8</td>
<td>154</td>
<td>98</td>
<td>100</td>
<td>0</td>
<td>X</td>
<td>rural</td>
</tr>
</tbody>
</table>

Note. Pop = Student Population; %F/RL = % Free and Reduced Lunch; %W = % White; %ELL = % English Language Learners.

The target independent variables for this study were focused on aspects of implementation and change: years of implementation, schools (third grade), and beginning year of implementation. Demographic and personal information matched at the level of the
individual student were available as control variables. Table 2 provides a descriptive comparison of the nine America's Choice Kentucky schools. The total school population along with the type and composition of each school are shown. The dependent variable was third grade mathematics NCE scores for respective years of America's Choice activity. All data files were taken from a secondary data base from the Kentucky Department of Education (KDE).

Population and Sample

Table 3

*America's Choice Components Implemented by Each School*

<table>
<thead>
<tr>
<th>School</th>
<th>Reading</th>
<th>Writing</th>
<th>Mathematics</th>
<th>Science</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>X</td>
<td>X</td>
<td>--</td>
<td>X</td>
</tr>
<tr>
<td>B</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>C</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>D</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>--</td>
</tr>
<tr>
<td>E</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>--</td>
</tr>
<tr>
<td>F</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>--</td>
</tr>
<tr>
<td>G</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>H</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>--</td>
</tr>
<tr>
<td>I</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>--</td>
</tr>
</tbody>
</table>

*Note.* X indicates component implemented; -- indicates component not implemented.

America's Choice provides the school with a plan for implementing the various content related components of the program. In the first year, schools implement the reading program followed by the writing program with an introduction to the mathematics component. During the second year the mathematics and science components are
implemented. Of the nine America’s Choice schools, eight implemented the mathematics component. Table 3 shows the elements of the America’s Choice design that were implemented by each of the schools.

Table 4

*Years of Implementation of America’s Choice Mathematics in Kentucky Schools*

<table>
<thead>
<tr>
<th>School</th>
<th>P1</th>
<th>Y1</th>
<th>Y2</th>
<th>C1</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>2003</td>
<td>2004</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>C</td>
<td>2001</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
</tr>
<tr>
<td>D</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
<td>2005*</td>
</tr>
<tr>
<td>E</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
<td>2005*</td>
</tr>
<tr>
<td>F</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
<td>2005*</td>
</tr>
<tr>
<td>G</td>
<td>2001</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
</tr>
<tr>
<td>H</td>
<td>2001</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
</tr>
<tr>
<td>I</td>
<td>2001</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
</tr>
</tbody>
</table>

*Note.* P = year prior to implementation; Y = year of implementation; C = continuation of America’s Choice in year after active implementation ended; N/A refers to schools that did not actively implement America’s Choice for various reasons. The mathematics component of America’s Choice was not implemented until the second year of the program (reading and writing in year 1). Thus the year prior to implementation for math was actually Year 1 of the project for reading and writing (compared to Table 1).

*This school did not actively implement America’s Choice during this year.*

Table 4 presents the years of implementation for third grade exiting primary students in Kentucky schools which implemented the mathematics component of the
research-based America’s Choice comprehensive design model beginning in 2002. Shown is the sequence for each school (prior year, years of active support, and continuation). School A did not participate in the mathematics component. School B was excluded from analyses because it did not participate in America’s Choice for Year 2 of implementation or the continuation year.

The population for the current investigation was defined consistent with the eight sampled schools in Table 4. Thus, the population was all third grade students in all Kentucky schools with CTBS/5 mathematics NCE scores for the years 2001 through 2005 ($N = 241,782$). This encompasses the year prior to America’s Choice mathematics as well as the continuation years, 2004 and 2005. Even though schools D-F did not actively continue America’s Choice subsequent to Year 2 of Implementation, the data for 2005 represent their next year, an equivalent time frame compared to schools C, G, H, and I.

**Effective Sample**

The effective sample for this study is the set of third grade exiting primary students collectively across the seven elementary schools implementing the mathematics component of the America’s Choice reform model for a minimum of two years. As noted above, the demographic variables in Table 2 are matched for each individual student. Table 5 shows the sample size for each of these schools (third grade) by year of implementation. The figure ($N = 1,443$) for the Total row and Total column represents the overall sample for schools C-I for years 2001-2005. These successive cohorts provide a window to the degree of school level sustained growth in mathematics at third grade over the years of implementation. These data are not longitudinal, however. The CATS accountability model in Kentucky compares successive cohorts, not the same students over time.
Table 5

*Number of Third Graders in Each School during Years of the Study*

<table>
<thead>
<tr>
<th>School</th>
<th>P1</th>
<th>Y1</th>
<th>Y2</th>
<th>C1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>50</td>
<td>39</td>
<td>42</td>
<td>48</td>
<td>179</td>
</tr>
<tr>
<td>D</td>
<td>43</td>
<td>48</td>
<td>61</td>
<td>50*</td>
<td>202</td>
</tr>
<tr>
<td>E</td>
<td>45</td>
<td>49</td>
<td>40</td>
<td>40*</td>
<td>174</td>
</tr>
<tr>
<td>F</td>
<td>105</td>
<td>112</td>
<td>100</td>
<td>103*</td>
<td>420</td>
</tr>
<tr>
<td>G</td>
<td>78</td>
<td>58</td>
<td>51</td>
<td>64</td>
<td>251</td>
</tr>
<tr>
<td>H</td>
<td>45</td>
<td>32</td>
<td>42</td>
<td>36</td>
<td>155</td>
</tr>
<tr>
<td>I</td>
<td>15</td>
<td>20</td>
<td>15</td>
<td>12</td>
<td>62</td>
</tr>
<tr>
<td>Total</td>
<td>381</td>
<td>358</td>
<td>351</td>
<td>353</td>
<td>1,443</td>
</tr>
</tbody>
</table>

*aThis school did not actively implement America's Choice during this year.*

Each of the seven schools in the effective sample is described below. The thumbnail sketches provide a summary of these rural, predominately white low socio-economic schools.

School C is a one building rural school located in a small independent school district in central Kentucky with a population of 600 preschool through twelfth grade students. Within the school, the grades were divided into three clusters representing elementary (preschool through grade four), middle (grades five through eight), and high (grades nine through twelve). School C was a Title I funded school with a predominantly white (90%) student population.

School D is located in a small rural district in Kentucky with a population of 312 preschool through fifth grade students. The student population was predominately white with a small representation of English language learners (3%) and African-Americans...
School D is a low socioeconomic status (SES) school as demonstrated by 93% of the student population participating in the free/reduced lunch program. This school did not participate in the Title I program or receive Title I funding.

School E is located in the same small rural district as School D. The preschool through fifth grade school served 247 students consisting of 1% African-American and 4% English language learners. Participating in the free/reduced lunch program are 66% of the student population. School E did not receive Title I funding.

School F was also located in the same small rural district as School D and School E. School F was a preschool through fifth grade school with a student population of 642 students. The student population is predominately white with a small representation of English language learners (3%) and African-American (1%) students. Of the total student population 38% participated in the free/reduced lunch program. During the first year of implementation of the America's Choice comprehensive reform model, School F was named a Blue Ribbon School by the U.S. Department of Education for making significant progress in closing the achievement gap based on scores during the previous three year test cycle. Even with this improvement, their overall low scores demonstrated the need for more significant change and led them to adopt America's Choice.

School G was a kindergarten through fifth grade school serving 335 students who are predominately white, located in one of Kentucky's smallest districts. School G has low income students and receives Title I funding to support instruction. Of the student population, 95% participated in the free and reduced lunch program.

School H was a kindergarten through fifth grade school located in the same small district in Kentucky as School G. The student population of 213 students is predominately white (99%). Of the population 95% qualified for the free/reduced lunch
program. This school also received Title I funding to support educational practices that enhance learning of low achieving students.

School I was located in a rural district near a military base in central Kentucky. The school served 154 kindergarten through eighth grade students. The student population was predominately white with 98% participating in the free and reduced lunch program.

Data Analysis

Starting with the 1998-99 school year, Kentucky implemented the Commonwealth Accountability Testing System (CATS) to assess school performance which included student achievement and non-cognitive measures. CATS is given in the spring each year to students in grade 3 (exiting primary) through grade 12. Students at different grades take tests in different subject areas. The Comprehensive Test of Basic Skills (CTBS/5) is given to students in grades three, six, nine and twelve. CTBS/5 results are reported in Normal Curve Equivalent (NCE), scale scores, and National Percentiles.

For this study, the researcher analyzed NCE mathematics scores from the CTBS/5 annual assessment for third grade (exiting primary) students in Kentucky elementary schools implementing the America’s Choice comprehensive reform model. Scores were collected across four cohorts, beginning with the year prior to implementation, the two years during implementation, and the year after supportive implementation ended. The scores from the year prior to implementation provided a base line score for each of the schools. Because CATS is a cohort model, these data do not constitute longitudinal scores on the same students.

As part of the analysis process, the researcher calculated descriptive statistics for all variables and provided population parameters in chart form. A full Pearson $r$ correlation matrix for the NCE Mathematics scores and a set of demographic factors is also presented.
For the primary analyses of the study, multiple regression models were used to examine the relationship between mathematics achievement and a set of predictors (i.e., year of implementation, schools, and beginning year of implementation). The predictors in the change model are measured nominally; dummy contrasts are coded for these variables. These dummy contrasts complete the set of data utilized in the regression equations. Research Questions 2-5 required multiple regression models to examine the relationship between mathematics achievement and the change model predictors (i.e., year of implementation, schools, and beginning year of implementation) plus demographic controls. The Statistical Package for Social Sciences (SPSS) computer package was used for all analyses.

Research Questions

The six Research questions are repeated for the convenience of the reader in this section. For each, the specific analytic procedure is described.

For the seven America’s Choice schools for third grade mathematics:

1. To what degree do the America’s Choice schools differ from the statewide mean for math achievement at the beginning of the program (Prior Year)?

2. To what extent is the year of implementation related to math achievement?

3. To what extent are differences between schools related to math achievement?

4. To what extent is the beginning year of the program related to math achievement?

5. What is the effect of controlling for demographic factors for Research Questions 2-4?

6. For the three years of the America’s Choice mathematics program (Year 1, Year 2, Continuation 1), how does the progress of the America’s Choice schools compare to that statewide for the same period?

The six empirical research questions required different types of analyses. These are
specified in turn.

For Research Question 1, the beginning of the program is defined as the year prior to implementation. The math means for each of the America’s Choice schools individually and as cohorts were compared to the state. Because of the staggered start, the mean for the statewide population for CTBS/5 grade 3 NCE mathematics scores is calculated across the two years as an average. For example, when calculating Prior Year, 2001 scores for schools C, G, H, and I were averaged with 2002 scores for schools D-F. To compare the seven America’s Choice schools to the statewide population with respect to their average math achievement for Year 1, Year 2, and Continuation Year, one sample t tests were conducted at $\alpha = .05$.

Research Questions 2-4 represent the effect of the America’s Choice program due to the three dimensions of implementation: the change over years of implementation, differences across schools, and effect of different beginning years of implementation. Research Question 5 provides evidence on the effect of controlling for demographic factors. For this study Research Questions 2-5 are all addressed via hierarchical regression. For Research Questions 2-4, a single hierarchical regression is computed with dummy codes representing change over years of implementation (entered as Step 1), differences across schools (entered as Step 2), and beginning year of implementation (entered as Step 3), respectively (see *Dummy Coding*, below). The steps of the hierarchical regressions are ordered based on the three components of the change model as reflected in Research Questions 2-4, respectively. In RQ5 a second hierarchical regression is calculated with the demographic factors added in Step 1, the implementation factors in Step 2, differences across schools in Step 3, and beginning year of implementation in Step 4. For both hierarchical regressions, the sequential entry displays the separate and additive effort of
each set of variables. Within the hierarchical steps, the relevant variables are entered
simultaneously.

Finally, Research Question 6 addressed the question of comparative change for the
seven America’s Choice schools versus the state population. This required two approaches.
First, one-sample \( t \) tests comparing third grade CTBS/5 mathematics mean NCE scores for
the America’s Choice schools to the State were computed, separately for the two cohorts
beginning in 2001 (Schools C, G-I) and 2002 (Schools D-F). Second, percentage change
scores (%Change) were calculated based on these third grade math scores, for both the
seven America’s Choice schools and the state. Percentage change scores are calculated as
the target year minus the prior year divided by the prior year. These change scores are
computed for three periods: (a) Year 1 minus Prior Year, (b) Year 2 minus Prior Year, and
(c) Continuation Year minus Prior Year.

Because of the staggered start date for implementing America’s Choice, the
calculations for the percentage change scores require the overall sample means across
schools C, G, H, and I (beginning 2001) and similarly for schools D, E, and F (beginning
year 2002): for Prior Year to implementation, for Year 1 of implementation, for Year 2 of
implementation, and for the Continuation Year, all as defined operationally below (see
Dummy Coding and Description of the Variables).

For the statewide calculations of percentage change scores for CTBS/5 third grade
NCE Mathematics, the population means that correspond to Schools C, G, H, and I are for

The next step is to produce one set of weighted percentage change scores (Year 1,
Year 2, Continuation Year) for the seven America’s Choice schools and one for statewide.
This is calculated by the following formula: (%Change across Schools C, G-I)(N for
Schools C, G-I) + (\%Change across Schools D-F)(N for Schools D-F) / (N for Schools C, G-I) + (N for Schools D-F). A parallel formula is used to calculate the weighted % change score for the state, with the Ns representing the statewide total population for 2001 and 2002, respectively.

Finally, the weighted percentage change scores for the entire sample (seven America’s Choice schools) are compared to the weighted percentage change statewide scores for the population (across 2001 and 2002). This is a descriptive analysis, with straightforward examination of the America’s Choice and State %Change scores for the two years of active implementation and the continuation year when support is no longer available. This comparative table is supplemented with a graph representing rate of progress across the years of the America’s choice intervention.

Multiple regression models are used in data analysis to study the relationships between the independent variables and dependent variables. There are several different approaches to conducting multiple regression analyses. For this study the researcher used a hierarchical regression model. For Research Questions 2-4, a single regression is conducted with the blocks of variables representing implementation change over years, school to school differences, and the staggered start date of America’s Choice program (see Dummy Coding below) entered successively in Steps 1-3. This gives an estimate of the separate effect of each block of variables on the dependent variable. For Research Question 5, a second hierarchical regression analysis is required. Again blocks of variables are entered in sequential steps that reflect a priori ordering of the effects of those variables (Tabachnick & Fidell, 2001). For this study, the demographic factors were entered in Step 1, giving the effect of those variables on the math outcomes. Then in Steps 2-4, the blocks of variables related to the implementation of America’s Choice are entered to determine
their impact, net of the demographic influence from Step 1: change over time in Step 2, school differences in Step 3, beginning year in Step 4. The analysis model is tested as to whether each successive block of target variables as a whole significantly increases $R^2$, given the demographic variables already entered into the regression equation.

**Assumptions of Regression Models**

Principal assumptions which justify the use of linear regression models include (a) linearity of the relationship between dependent and independent variables; (b) normality of the error distribution; (c) independence of the predictors; and (d) homoscedasticity of the errors. If any of these assumptions is violated, the forecasts yielded by a regression model can be seriously biased or misleading. Therefore, prior to analyzing data using multiple regression models, several statistical techniques were employed to check the assumptions of a regression analysis. Such techniques focus primarily on analyzing residuals, assessing the influence of outliers, and assessing collinearity.

Residual plots were examined to check the linearity, normality, and homoscedasticity assumptions. Linearity is usually most evident in a plot of residuals versus predicted values, which is a part of standard regression output. The points should be symmetrically distributed around a horizontal line in the plot of residuals versus predicted values. Examination of the residuals is also useful to detect violation of the homoscedasticity assumption, in that it can be an alert for evidence of residuals that are changing (i.e., more or less spread out) as a function of the predicted value.

Normality was checked through a normal probability plot of the residuals, which compares a plot of the error distribution versus a normal distribution having the same mean and variance. If the distribution is normal, the points on this plot should fall close to the diagonal line. Another useful approach to detecting nonnormality is to observe outliers. In
linear regression an outlier is an observation with a large residual or an unusual dependent variable value (Allison, 1999). An outlier may indicate a sample peculiarity or may indicate a data entry error or other problem. A scatter plot was used to assess for the influence of outliers in this study.

Predictors that are highly collinear can cause problems in estimating the regression coefficients. One problem is an increase in the standard error of the coefficients, which makes it difficult to get an accurate measure of the regression coefficients of those collinear variables (Allison, 1999). The large standard error can cause small differences to be magnified and the regression slope to be unstable because of the mathematics of the regression model. The unstable regression slope can produce a large, significant $R^2$ without any of the coefficients being significant. To check for degree of collinearity the researcher used the variance inflation factor or VIF. Tolerance, defined as $1/VIF$ (variance inflation factor) was used to check on the degree of collinearity. A tolerance value lower than 0.1 is comparable to a VIF of 10. For this study, all VIF values were less than 10 so we can assume the multicollinearity problem did not occur.

Model Specifications

Research Questions 2-5 are based on a hierarchical regression model hypothesizing that students’ math achievement scores can be predicted by a linear combination of year of implementation of the America’s Choice program, schools, and beginning year of implementation as target variables, plus a set of school demographics as control factors. With hierarchical regression, these different blocks are entered sequentially. An example of this prediction model can be specified for Step 1 of RQ5 as follows:

$$MATH = \beta_0 + \beta_1(SIZE) + \beta_2(%F/RL) + E \quad (1)$$

where MATH = third grade NCE math scores on comprehensive Test of Basic Skills
(CTBS/5); SIZE = school size; %F/RL = percentage of students in each school participating in the free and reduced lunch program; and E is the error component. Within this first step of the overall hierarchical regression, these variables are considered simultaneously.

**Dummy Coding**

This study includes two categorical variables with more than two levels on each. Since multi-level categorical predictor variables cannot be entered directly into a regression model and be meaningfully interpreted, additional steps are needed. These include recoding the categorical variables into a number of separate dichotomous variables, called dummy coding. Dichotomous variables have the advantage that they can be directly entered into the regression model. This dummy coding scheme is used for three categorical variables relating to the implementation of America’s Choice: (a) year of implementation (YR), (b) schools (SCH), and (c) beginning year of implementation (BGYR).

Using dummy coding, a categorical variable with $k$ levels is transformed into $k-1$ variables each with two levels. For example, the original SCH variable with seven levels is converted to six dichotomous variables that contain the same information as the single variable. Setting the first level on the predictor (i.e., school C) as a reference value in the coding scheme, the six dummy variables are created for indexing the seven categories of the SCH variable. The other two categorical predictors YR and BGYR are transformed in the same way. With additional terms as defined by dummy coding above, the regression model specified in Equation 1 can be expressed to represent the sequential steps of the single hierarchical regression for RQs 2-4. For example, Step 1 (the calculations for RQ2) would take the following form:

$$ MATH = \beta_0 + \beta_1(Z'_1) + \beta_2(Z'_2) + \beta_3(Z'_3) + E $$

(2)
The six school \( Z \) contrasts would be added in Step 2 (calculation for RQ3); the beginning year of implementation \( Z^* \) contrast would be added in like fashion in Step 3 (RQ4 computation). The various \( Z \) contrasts are coded as follows:

\[
Z_1 = \begin{cases} 
1 & \text{if } SCH = \text{School D} \\
0 & \text{otherwise} 
\end{cases}
\]

\[
Z_2 = \begin{cases} 
1 & \text{if } SCH = \text{School E} \\
0 & \text{otherwise} 
\end{cases}
\]

\[
Z_3 = \begin{cases} 
1 & \text{if } SCH = \text{School F} \\
0 & \text{otherwise} 
\end{cases}
\]

\[
Z_4 = \begin{cases} 
1 & \text{if } SCH = \text{School G} \\
0 & \text{otherwise} 
\end{cases}
\]

\[
Z_5 = \begin{cases} 
1 & \text{if } SCH = \text{School H} \\
0 & \text{otherwise} 
\end{cases}
\]

\[
Z_6 = \begin{cases} 
1 & \text{if } SCH = \text{School I} \\
0 & \text{otherwise} 
\end{cases}
\]

for the school (SCH) variable,

\[
Z'_1 = \begin{cases} 
1 & \text{if } YR = 1 \text{ (2002 for Schools C, G - I; 2003 for Schools D - F)} \\
0 & \text{otherwise} 
\end{cases}
\]

\[
Z'_2 = \begin{cases} 
1 & \text{if } YR = 2 \text{ (2003 for Schools C, G - I; 2004 for Schools D - F)} \\
0 & \text{otherwise} 
\end{cases}
\]

\[
Z'_3 = \begin{cases} 
1 & \text{if } YR = 3 \text{ (2004 for Schools C, G - I; 2005 for Schools D - F)} \\
0 & \text{otherwise} 
\end{cases}
\]

for the year of implementation (YR) variable, and

\[
Z^* = \begin{cases} 
1 & \text{if } BGYR = 2003 \text{ (Year 1 for Schools D - F)} \\
0 & \text{if } BGYR = 2002 \text{ (Year 1 for Schools C, G - I)} 
\end{cases}
\]

for the beginning year of implementation (BGYR) variable (see Table 4).
Description of the Variables

For the purpose of this study the researcher included two types of independent variables, change and control. The dependent variables are the third grade Comprehensive Test of Basic Skills (CTBS/5) NCE mathematics scores. Detailed information about each type of variable is outlined in the following sections.

Independent Variables

America’s Choice represents a specific model of comprehensive reform, an intervention program designed to bring about change. For this study, the degree of change is reflected in the dependent variables. The model of change represents the factors that can be shown to be associated with any progress made during the implementation of America’s Choice: (a) change due to implementation for a given year of the program, (b) differences in change from one school to the next, and (c) different starting year for the schools which is also of interest as statewide accountability scores in Kentucky have improved significantly under the KERA reforms (Poggio, 2000). In several studies of Kentucky Scholastic Audits, year of implementation was not only significant but had an influence greater than the free/reduced lunch status (Ennis, 2007; McKinney, 2007; Saravia, 2008).

Implementation Model

Schools (SCH): Data for seven Kentucky elementary schools implementing the mathematics component of the America’s Choice comprehensive school reform model for a minimum of two years were included in this study. Schools C, G, H, and I include students from four successive cohorts across 2001-2004. Schools D, E, and F include the four cohorts for 2002-2005. The original SCH variable is dummy coded in the data set for the regression analyses (see Dummy Coding, above).

Year of Implementation (YR): Year of Implementation includes the year prior to
implementation, two years of implementation, and a continuation year. The nominal scale
with four levels is Dummy coded with prior year of implementation as the referent
category, with three Z contrasts: \( Z_1 = \) first year of implementation; \( Z_2 = \) second year of
implementation; \( Z_3 = \) continuation year of implementation. Thus, the YR variable with the
four categories is transformed into three separate dichotomous variables (see Dummy
Coding, above).

**Beginning Year of Implementation (BGYR):** The year that each of the seven
schools began implementation of America’s Choice mathematics was either 2002 or 2003.
To account for this staggered start, BGYR is conceptualized as a nominal variable
measured at the level of the individual. Each student in schools C, G, H, and I (2002 start
date) is coded as 0; each student in schools D-F (2003 start date) is coded as 1.

**Control Variables**

The control variables for this study include the demographic factors (taken from
Table 2). Type of school (all rural for the seven schools) is excluded from analyses.
Likewise, Title I status is excluded. By chance (not planned), schools D-F were the three
non-Title I sites. This is co-terminus with Beginning Year of Implementation so that any
variance due to Title I status would already be captured by the beginning year variable.

Percentage white (%W) and percentage English language learners (%ELL) were
excluded from the analysis of demographic variables. Table 2 showed that the seven
schools implementing the mathematics component of America’s Choice were
predominately white and therefore demonstrated little variation. Likewise, the percentage
of English language learners for the schools included in this study was so restricted that the
inclusion of this variable would not produce significant results as part of the analysis (see
Table 2).
Size (SIZE): Total number of third grade exiting primary students cumulative across four cohorts for each of the seven schools, coded as a ratio scale (see Table 5).

Percentage Free/Reduced Lunch (%F/RL): The percentage of students in each school who have been identified as participating in the free and reduced lunch program, this ratio is a proxy for socioeconomic status (SES), coded as 1 = participation in the program, and 0 = not participating. Although the measure is problematic (cf. Harwell & LeBeau, 2010), it is the most influential demographic marker in most studies of school achievement.

Dependent Variable

The dependent variable for this study is based on the set of Comprehensive Test of Basic Skills (CTBS/5) mathematics NCE scores for third grade exiting primary students in each of the Kentucky schools implementing the mathematics component of America's Choice. The CTBS/5 is a multiple choice, nationally standardized achievement test that compares student's performance to a national norm group. The questions address the components of number/operations, algebra, data analysis, geometry, and measurement.

This study included the scores for the baseline year, two years of implementation, and year of continuation, which represents the entire set of the Comprehensive Test of Basic Skills (CTBS/5) mathematics NCE scores. However, three schools did not collect data for the continuation year as part of the project per se. Although the researcher was able to attain these data from statewide files, it is unknown whether the continuation scores for the subgroups of schools would be different. Those who were still “doing” America’s Choice (still cognizant of and still working to implement) may be different from (higher than) those scores that coincided with the next year of life in schools for whom America’s Choice was officially over (presumably less attention to/concern with implementation).
Given this caveat, the dependent variable is measured as the NCE scores for CTBS/5 third grade NCE mathematics achievement across the four data collection points from prior year to the year after the implementation of the America’s Choice program.

Validity and Reliability

Validity and reliability are two important concepts for any research study. Validity refers to the measurement instrument’s ability to measure accurately what it is supposed to measure free of systematic error (Vogt, 1999). Reliability is the consistency or stability of a measure or test from one use to the next. The measurement instrument is said to be reliable when repeated measurements produce identical or similar results (Vogt).

The Kentucky Department of Education (KDE) contracted with an independent firm, the Human Resources Research Organization (HumRRO), to conduct validity and reliability research for the Commonwealth Accountability Testing System (CATS). The annual reports HumRRO produced contain statistical data, tables, and information relating to the validity and reliability of the Comprehensive Test of Basic Skills (CTBS/5) which is the third grade portion of the CATS assessment system for Kentucky. In addition, the Kentucky Department of Education is in contact with the National Technical Advisory Panel for Assessment and Accountability (NTAPAA) which is a panel of assessment experts that advises the Kentucky Legislature and the KDE on assessment related matters.

The report produced by HumRRO examined the stability of the CTBS scores and other aspects of the CATS assessment across grades. Results indicate that students who do well on CTBS one year are likely to do well on CTBS on subsequent years. Correlations between the same content areas (for Reading and Math) across grades ranged $r = .62$ to $r = .73$. Analyses were also conducted to examine the stability of demographic differences in scores over time. The results indicate that the gender, socioeconomic, and racial
differences in scores remained relatively constant over time for CTBS. The information from this report provides adequate evidence for the validity and reliability of CTBS scores and demographic factors related to student achievement.

Ethical Considerations

Ethical research assumes that all humans participating in a research study have the right to be protected from any harm. This study was presented for review by the Institutional Review Boards of University of Louisville and Western Kentucky University to ensure the confidentiality and protection of the participants. Approval documentation from both institutions is included in Appendix B and C. The data used in this study were collected through a third party, the Kentucky Department of Education, and not directly by the researcher. All information has been coded so no individual schools or students are identified in the study, thus minimizing threat to individual human subjects.

Summary

This study examined the impact of implementing the mathematics component of America’s Choice comprehensive reform for third grade exiting primary students in seven Kentucky schools. The Normal Curve Equivalent (NCE) mathematics scores from the Comprehensive Test of Basic Skills (CTBS/5) data were collected and analyzed for the year prior to implementation, the two years during implementation, and the year after supported implementation to determine the effect of the mathematics component.

The researcher used quantitative analyses of the data collected from the Comprehensive Test of Basic Skills (CTBS/5) during a five year period for the schools in the study. For this study the independent variables included change variables (schools, years of implementation of the mathematics component, and beginning year of implementation) and control variables (demographic factors). The dependent variable was
Normal Curve Equivalency (NCE) mathematics scores for third grade students.

Purposeful sampling was used to select the Kentucky schools that participated in this study. The seven Kentucky elementary schools included in the study met the criteria by selecting and implementing the America's Choice comprehensive school reform model, specifically, the mathematics component. Research Question 1 utilized one sample t tests comparing scores of sample students (the seven schools) to the population scores (statewide). Research Questions 2-4 utilized a single hierarchical regression with separate blocks of variables entered in Steps 1-3 representing RQs 2-4, respectively. Research Question 5 also required hierarchical regression; however, here the demographic factors were entered in Step 1 to determine the prior effects of these control variables. Research Question 6 was examined in two ways. First, one sample t tests compared the progress of the America's Choice schools to the state as a whole for Year 1 of implementation, Year 2, and the continuation year after active support ended. In the second approach, percentage change (%Change) scores were calculated for both the sample schools and the state population. These percentage change scores were then compared to the state using chi-square analysis for Year 1, Year 2, and the Continuation year, producing a 3 x 2 matrix. For all three years, the %Change was computed by comparing the target year to the baseline (prior year of data).

Validity and reliability data on the assessment instrument for this study was conducted by the Kentucky Department of Education through two sources, the Human Resources Research Organization (HumRRO) and the National Technical Advisory Panel for Assessment and Accountability (NTAPAA).
CHAPTER IV

RESULTS

Introduction

The purpose of this quantitative study was to add to the knowledge base of research regarding the effectiveness of America’s Choice comprehensive reform. Specifically, this research examined the effects of implementing the mathematics component for third grade exiting primary students in seven Kentucky elementary schools. This builds upon the work of VanMeter (2005) who studied the impact of the reading component of America’s Choice in the same set of Kentucky elementary schools.

The Comprehensive Test of Basic Skills (CTBS/5) Normal Curve Equivalent (NCE) mathematics scores were collected across four cohorts that represented the year prior to implementation (P1), two years of supported implementation (Y1 and Y2), and the continuation year with no support (C1). Because of staggered start dates, schools D-F used 2003 data and schools C, G, H, and I used 2002 data to represent Year 1 of Implementation. The central research question guiding this study was, What are the effects of implementing America’s Choice comprehensive school reform design in mathematics in seven elementary schools in Kentucky?

This chapter contains the results of the quantitative analyses utilizing simultaneous and hierarchical regressions to examine the relationships of the variables for each of the research questions guiding the study. The Statistical Package for the Social Sciences (SPSS) computer program was used for all analyses. The research
questions guiding this study are repeated below for the reader’s convenience.

For the seven America’s Choice schools for third grade mathematics:

1. To what degree do the America’s Choice schools differ from the statewide mean for math achievement at the beginning of the program (Prior Year)?

2. To what extent is the year of implementation related to math achievement?

3. To what extent are differences between schools related to math achievement?

4. To what extent is the beginning year of the program related to math achievement?

5. What is the effect of controlling for demographic factors for Research Questions 2-4?

6. For the three years of the America’s Choice mathematics program (Year 1, Year 2, Continuation 1), how does the progress of the America’s Choice schools compare to that statewide for the same period?

The remaining sections of this chapter address the descriptive statistics for the variables, the results of analyses for each research question, and a summary of the results by school. The chapter concludes with a summary.

Descriptive Statistics

This study analyzed mathematics scores for third grade exiting primary students in seven Kentucky elementary schools, representing part of Kentucky’s accountability formula, as obtained from the Kentucky Department of Education. This research reported descriptive statistics for the independent variables: control (demographic variables—size and percent free/reduced lunch), and implementation model (school, years of implementation, and beginning year of implementation). The dependent variable was the Comprehensive Test of Basic Skills (CTBS/5) Normal Curve Equivalent (NCE) mathematics scores.

Independent Variables
The study included two types of independent variables, control and implementation model. The control variables were the demographic factors: Size, number of grade 3 students across four years of America’s choice (SIZE); and Percent Free/Reduced Lunch (%F/RL).

Table 6 presents the descriptive statistics for the demographic variables size and percent free/reduced lunch (%F/RL) for each of the seven schools. Since the study only focused on third grade students, the number of students in each cohort is relatively low, as reflected in Mean, Minimum, and Maximum values for each school across the four cohorts. School I is the smallest ($M = 15.5$) while School F is the largest ($M = 104.5$).

For third grade students classified as participating in the free/reduced lunch program, School G ($M = 69.5\%$) was the highest, while School F ($M = 12.0\%$) was the lowest. These figures, however, are considerably lower than the schoolwide participation in the federal free/reduced lunch program for the year 2000 (see Table 2 in Chapter III) where School F was the lowest at 38% and School I was the highest at 98%. For Table 2 (all of the students in the school for the year 2000), only two schools (E and F at 66% and 38%, respectively) had percent free/reduced lunch less than 90. In contrast, for third grade only across the four years of the study (2001-2004 or 2002-2005 for Table 6), only schools G, H, and I had percentages above 50.
Table 6

*Descriptive Statistics for Demographic Variables Size and Percent Free/Reduced Lunch*

*for Third Grade across the Seven Schools*

<table>
<thead>
<tr>
<th>School</th>
<th>$N$</th>
<th>$M$</th>
<th>$SD$</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SIZE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>179</td>
<td>45.3</td>
<td>5.18</td>
<td>39</td>
<td>50</td>
<td>11</td>
</tr>
<tr>
<td>D</td>
<td>202</td>
<td>50.8</td>
<td>7.58</td>
<td>43</td>
<td>61</td>
<td>18</td>
</tr>
<tr>
<td>E</td>
<td>174</td>
<td>44.0</td>
<td>3.82</td>
<td>41</td>
<td>49</td>
<td>8</td>
</tr>
<tr>
<td>F</td>
<td>420</td>
<td>104.5</td>
<td>5.20</td>
<td>100</td>
<td>112</td>
<td>12</td>
</tr>
<tr>
<td>G</td>
<td>251</td>
<td>63.0</td>
<td>11.13</td>
<td>52</td>
<td>78</td>
<td>26</td>
</tr>
<tr>
<td>H</td>
<td>155</td>
<td>38.3</td>
<td>5.18</td>
<td>32</td>
<td>45</td>
<td>13</td>
</tr>
<tr>
<td>I</td>
<td>62</td>
<td>15.5</td>
<td>3.31</td>
<td>12</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>1443</td>
<td>51.61</td>
<td>5.91</td>
<td>12</td>
<td>112</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th>%F/RL</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>179</td>
<td>42.8</td>
<td>6.25</td>
<td>36</td>
<td>49</td>
<td>13</td>
</tr>
<tr>
<td>D</td>
<td>202</td>
<td>40.8</td>
<td>8.01</td>
<td>27</td>
<td>47</td>
<td>20</td>
</tr>
<tr>
<td>E</td>
<td>174</td>
<td>27.3</td>
<td>2.05</td>
<td>24</td>
<td>29</td>
<td>5</td>
</tr>
<tr>
<td>F</td>
<td>420</td>
<td>12.0</td>
<td>5.87</td>
<td>7</td>
<td>22</td>
<td>15</td>
</tr>
<tr>
<td>G</td>
<td>251</td>
<td>69.5</td>
<td>2.69</td>
<td>65</td>
<td>71</td>
<td>6</td>
</tr>
<tr>
<td>H</td>
<td>155</td>
<td>65.5</td>
<td>5.17</td>
<td>58</td>
<td>72</td>
<td>14</td>
</tr>
<tr>
<td>I*</td>
<td>50</td>
<td>79.0</td>
<td>14.89</td>
<td>67</td>
<td>100</td>
<td>33</td>
</tr>
<tr>
<td>Total</td>
<td>1431</td>
<td>45.14</td>
<td>6.42</td>
<td>7</td>
<td>100</td>
<td>93</td>
</tr>
</tbody>
</table>

*Note.* SIZE = grade 3 student population across four years; %F/RL = percent Free/Reduced Lunch.

*aSchool I did not report any Free and Reduced Lunch data for 2004.
The implementation model variables were schools (SCH), Year of Implementation (YR), and Beginning Year of Implementation (BGYR). For this study, the implementation variables were coded nominally. Dummy Coding and Z contrasts were utilized to enter the variables into the regression model. The Dummy Coding produced separate dichotomous variables as follows (see Dummy Coding in Chapter III):

- $Z_1-Z_6$ (for schools D-I with School C as the referent category; Year of Implementation entered as $Z'_1$ = first year of implementation, $Z'_2$ = second year of implementation, and $Z'_3$ = continuation year with prior year as the referent category; and $Z''$ coded 1 = beginning year 2003 for schools D-F, 0 = beginning year 2002 for schools C, G-I. No descriptive tables are given for these dummy codes.

A Pearson $r$ correlation matrix for all continuous variables in the study is located in Appendix E. This includes the two demographic controls and the dependent variable, the NCE math scores. The calculations are based on the total number of students in each school across the four cohorts of third graders. Multiple regressions to answer the research questions are based on this matrix. Correlations among the variables are not very strong. The highest correlation was $r = .140$ for size with NCE mathematics scores. This demonstrates a weak, positive correlation, not significant, between these two factors.

**Dependent Variable**

The dependent variable for this study was the Comprehensive Test of Basic Skills (CTBS/5) Normal Curve Equivalent (NCE) mathematics scores for third grade exiting primary students in seven Kentucky elementary schools implementing America's Choice comprehensive reform. The scores collected represent the year prior to implementation of the mathematics component (P1), two years of supported implementation (Y1 and Y2), and the continuation year (C1) with no supported implementation. Due to the staggered
start date of the seven schools (2002 for schools C, G-I and 2003 for schools D-F) data were collected over a five year period to capture the appropriate data points.

Table 7 lists the mathematics scores for the seven America’s Choice schools across the four years of implementation. The overall total now indicates not much change in math achievement scores over the four years, with a slight increase for Year 2 ($M = 54.1$); however, the Continuation Year ($M = 50.1$) falls to a level below that at the beginning of the program ($M = 51.1$). School C was the only one to demonstrate an increase from Prior Year ($M = 49.4$) through Continuation Year ($M = 55.6$). School D experienced little change during implementation years (P1, $M = 52.8$ vs. Y2, $M = 52.3$) but dropped during the Continuation Year ($M = 46.3$). Schools also experienced little or no change during the active years of implementation of the America’s Choice mathematics component and regressed during the Continuation Year ($M = 43.4$). The math mean scores for School E increased from Prior Year of Implementation ($M = 50.3$) through the second year of implementation ($M = 56.0$) but fell during the Continuation year ($M = 51.8$). Schools F and H remained steady during the years of implementation; however, School F showed a decline for Continuation Year ($M = 53.7$).
Table 7

NCE Grade 3 Mathematics Scores for Seven America's Choice Schools across the Four Implementation Years

<table>
<thead>
<tr>
<th>School</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>N</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>50</td>
<td>49.4</td>
<td>17.5</td>
<td>42</td>
<td>51.6</td>
<td>21.8</td>
<td>48</td>
<td>55.6</td>
<td>20.7</td>
<td>179</td>
<td>52.0</td>
<td>20.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>43</td>
<td>52.8</td>
<td>19.8</td>
<td>61</td>
<td>51.2</td>
<td>19.4</td>
<td>50</td>
<td>46.3</td>
<td>18.8</td>
<td>202</td>
<td>50.6</td>
<td>17.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>45</td>
<td>50.3</td>
<td>17.3</td>
<td>40</td>
<td>56.0</td>
<td>15.1</td>
<td>40</td>
<td>51.8</td>
<td>16.4</td>
<td>174</td>
<td>53.6</td>
<td>16.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>105</td>
<td>61.7</td>
<td>18.5</td>
<td>100</td>
<td>60.3</td>
<td>22.2</td>
<td>103</td>
<td>53.7</td>
<td>19.2</td>
<td>420</td>
<td>58.0</td>
<td>20.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>78</td>
<td>46.9</td>
<td>20.8</td>
<td>51</td>
<td>61.6</td>
<td>24.7</td>
<td>64</td>
<td>48.6</td>
<td>19.1</td>
<td>251</td>
<td>51.4</td>
<td>21.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>45</td>
<td>50.6</td>
<td>14.3</td>
<td>42</td>
<td>50.3</td>
<td>20.8</td>
<td>36</td>
<td>51.1</td>
<td>23.2</td>
<td>155</td>
<td>50.7</td>
<td>19.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>15</td>
<td>46.0</td>
<td>18.9</td>
<td>15</td>
<td>45.9</td>
<td>17.9</td>
<td>12</td>
<td>39.6</td>
<td>18.5</td>
<td>62</td>
<td>43.8</td>
<td>17.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>381</td>
<td>52.8</td>
<td>18.2</td>
<td>358</td>
<td>53.2</td>
<td>19.0</td>
<td>351</td>
<td>55.6</td>
<td>20.1</td>
<td>353</td>
<td>51.0</td>
<td>19.4</td>
<td>1443</td>
<td>53.2</td>
<td>19.2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. P1 = Prior Year, Y1 = Year 1 of Implementation, Y2 = Year 2 of Implementation, CI = Continuation Year with no support.

For schools C, G-I, P1 of four years was 2001.

For schools D-F, P1 of four years was 2002.
Variable Coding

Table 8

Variables, Variable Label Codes, and Type of Data Utilized

<table>
<thead>
<tr>
<th>Variable</th>
<th>Variable Label Code</th>
<th>Type of Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>SIZE</td>
<td>Ratio Scale</td>
</tr>
<tr>
<td>Percentage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Free/Reduced Lunch</td>
<td>%F/RL</td>
<td>Ratio Scale</td>
</tr>
<tr>
<td>Implementation Models</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schools</td>
<td>SCH</td>
<td>Nominal Scale</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Dummy Codes, Z₁-Z₆)</td>
</tr>
<tr>
<td>Year of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implementation</td>
<td>YR</td>
<td>Nominal Scale</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Dummy Codes, Z₁₁-Z₃₃)</td>
</tr>
<tr>
<td>Beginning Year of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implementation</td>
<td>BGYR</td>
<td>Nominal Scale</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Dummy Code Z&quot;)</td>
</tr>
<tr>
<td>Dependent Variable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCE Mathematics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scores</td>
<td>NCE</td>
<td>Interval Scale</td>
</tr>
</tbody>
</table>

Table 8 presents a summary of the blocks of variables, variable codes, and the type of measurement used in the analysis process. These constitute the full data set and are the basis of subsequent hierarchical multiple regressions and one sample \( t \) tests that
address the research questions.

Research Questions

This section addresses the analysis of the six research questions that guided the research. Each question is listed in turn prior to the results.

Research Question 1

To what degree do the America's Choice schools differ from the statewide mean for math achievement at the beginning of the program (Prior Year)?

Because of the staggered start dates, Prior Year (P1) was 2001 for schools C, G, H, and I and 2002 for schools D-F. For Research Question 1, the Comprehensive Test of Basic Skills (CTBS/5) Normal Curve Equivalent (NCE) mathematics mean scores for each of the seven schools were compared to the statewide mean using one sample t tests which determine if sample scores are significantly different from the performance of the whole population, conducted at \( \alpha = .05 \).

Table 9 shows the results of the one sample t tests for each of the seven schools and for the schools in the two cohorts for the year prior to implementation (Prior Year) compared to the state mean. Schools C, G-I were compared to the statewide mean for 2001 (\( \mu = 54 \)) while Schools D-F were compared to the statewide mean for 2002 (\( \mu = 55.1 \)).

The findings for Prior Year show that School F, \( t(102) = 3.63, p < .001 \), was achieving above the state mean and School G, \( t(77) = -3.02, p = .003 \), below the state, both significantly different. The results for Schools C, G-I combined, \( t(185) = -4.21, p < .001 \) demonstrated statistical significance when compared to the state, while the results for Schools D-F combined were slightly above the state but not significant, \( t(190) = 1.40, p = .16 \). The overall result for Schools D-F masks the fact that School F (\( M = 61.7 \))
was significantly above the state mean while Schools D and E were below the mean, but not significantly so.

Table 9

*One Sample t Tests Comparing Means for the Seven Schools to the State Mean for Prior Year*

<table>
<thead>
<tr>
<th>School</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>t</th>
<th>p</th>
<th>95% Confidence Lower</th>
<th>95% Confidence Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>50</td>
<td>49.4</td>
<td>17.5</td>
<td>-1.88</td>
<td>.06</td>
<td>-9.60</td>
<td>.32</td>
</tr>
<tr>
<td>D</td>
<td>43</td>
<td>52.8</td>
<td>19.8</td>
<td>-.76</td>
<td>.45</td>
<td>-8.47</td>
<td>3.84</td>
</tr>
<tr>
<td>E</td>
<td>45</td>
<td>50.3</td>
<td>17.26</td>
<td>-1.85</td>
<td>.07</td>
<td>-9.95</td>
<td>1.42</td>
</tr>
<tr>
<td>F</td>
<td>105</td>
<td>61.7</td>
<td>18.52</td>
<td>3.63</td>
<td>&lt;.01</td>
<td>3.01</td>
<td>10.25</td>
</tr>
<tr>
<td>G</td>
<td>78</td>
<td>46.9</td>
<td>20.8</td>
<td>-3.02</td>
<td>&lt;.01</td>
<td>-11.79</td>
<td>-2.42</td>
</tr>
<tr>
<td>H</td>
<td>45</td>
<td>50.6</td>
<td>14.3</td>
<td>-1.57</td>
<td>.12</td>
<td>-7.82</td>
<td>.98</td>
</tr>
<tr>
<td>I</td>
<td>15</td>
<td>46.0</td>
<td>18.9</td>
<td>-1.64</td>
<td>.13</td>
<td>-18.47</td>
<td>2.47</td>
</tr>
<tr>
<td>Total</td>
<td>188</td>
<td>58.6</td>
<td>17.88</td>
<td>-4.21</td>
<td>&lt;.01</td>
<td>-8.32</td>
<td>-3.01</td>
</tr>
</tbody>
</table>

*a* For schools C, G-I, P1 was 2001 (μ = 54).

*b* For schools D-F, P1 was 2002 (μ = 55.1).

*Research Question 2*

To what extent is the year of implementation related to math achievement?

Research Questions 2-4 utilize the results of a single hierarchical regression to address the effects of three dimensions on the NCE third grade math scores: change over years of implementation (RQ2), differences between schools (RQ3), and different
beginning years (RQ4). The predictor variables are based on nominal coding utilizing the following Z contrasts (as explained in Chapter III): schools D-I entered as $Z_1-Z_6$ with School C as the referent category; Year of Implementation entered as $Z_1' =$ first year of implementation ($Y_1$), $Z_2' =$ second year of implementation ($Y_2$), and $Z_3' =$ continuation year ($C_1$) with prior year ($P_1$) as the referent category; and $Z''$ coded 1 = beginning year 2003 for schools D-F, 0 = beginning year 2002 for schools C, G-I. The dependent variable was the Comprehensive Test of Basic Skills (CTBS/5) Normal Curve Equivalent (NCE) mathematics scores for third grade exiting primary students in each of the seven Kentucky elementary schools implementing the mathematics component of America’s Choice reform.

This hierarchical regression model consists of three sequential blocks with variables within each step entered simultaneously. The first block (Step 1) represents Research Question 2 (year of implementation, $Z_1'$, $Z_2'$, and $Z_3'$); the second block (Step 2) addresses Research Question 3 (differences between schools, $Z_1-Z_6$); and the third block (Step 3) examines Research Question 4 (beginning year of program, $Z''$). Table 10 presents the single hierarchical regression for the full model (RQs 2-4, combined). The variance in the within model degrees of freedom in the ANOVA equations compared to the overall sample of 1443 is due to casewise deletion caused by unreported data and changes in student population. After evaluating the multiple regression assumptions, all of the model equations were acceptable based on the collinearity and other diagnostic analyses.

For Research Question 2, the relevant portion of the Table 10 hierarchical regression is Step 1 (Year of Implementation). This question examined the impact of implementation year (Year 1, Year 2, and Continuation Year) on math achievement. The
model for Step 1 is significant $F(3, 1440) = 3.25, p = .02$ with Adjusted $R^2$ of .01 indicating virtually no effect for year of implementation on math achievement. The $Z_2'$ contrast for year of implementation is significant, but this is a function of the large sample size. Cohen, Cohen, West, and Aiken (2003) indicate that for large $N$, the change in Adjusted $R^2$ should be a minimum of .02 to be substantive. Thus the largest standardized beta of .07 for Year 2 ($Z_2'$), interpreted as a change of .07 standard deviation units on the third grade math NCE scores associated with a one standard deviation change for the $Z_2'$ contrast (Year 2 compared to Prior Year), is essentially no difference.
Table 10

Hierarchical Regression of Mathematics NCE Scores on Three Blocks of America's Choice Implementation Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>$B$</th>
<th>$SE B$</th>
<th>Beta</th>
<th>$t$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>52.74</td>
<td>1.02</td>
<td></td>
<td>51.53</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>$Z'_1$</td>
<td>.50</td>
<td>1.47</td>
<td>.01</td>
<td>.34</td>
<td>.74</td>
</tr>
<tr>
<td>$Z'_2$</td>
<td>3.03</td>
<td>1.47</td>
<td>.07</td>
<td>2.06</td>
<td>.04</td>
</tr>
<tr>
<td>$Z'_3$</td>
<td>-1.54</td>
<td>1.47</td>
<td>-.03</td>
<td>-1.05</td>
<td>.29</td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>52.10</td>
<td>1.69</td>
<td></td>
<td>30.92</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>$Z'_1$</td>
<td>.34</td>
<td>1.45</td>
<td>.01</td>
<td>.24</td>
<td>.81</td>
</tr>
<tr>
<td>$Z'_2$</td>
<td>3.06</td>
<td>1.46</td>
<td>.07</td>
<td>2.10</td>
<td>.04</td>
</tr>
<tr>
<td>$Z'_3$</td>
<td>-1.65</td>
<td>1.45</td>
<td>-.04</td>
<td>-1.14</td>
<td>.26</td>
</tr>
<tr>
<td>$Z_1$</td>
<td>-1.80</td>
<td>2.00</td>
<td>-.03</td>
<td>-.90</td>
<td>.37</td>
</tr>
<tr>
<td>$Z_2$</td>
<td>1.02</td>
<td>2.07</td>
<td>.02</td>
<td>.49</td>
<td>.62</td>
</tr>
<tr>
<td>$Z_3$</td>
<td>5.38</td>
<td>1.74</td>
<td>.12</td>
<td>3.09</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>$Z_4$</td>
<td>-1.05</td>
<td>1.91</td>
<td>-.02</td>
<td>-.55</td>
<td>.58</td>
</tr>
<tr>
<td>$Z_5$</td>
<td>-1.91</td>
<td>2.15</td>
<td>-.03</td>
<td>-.89</td>
<td>.38</td>
</tr>
<tr>
<td>$Z_6$</td>
<td>-8.11</td>
<td>2.88</td>
<td>-.08</td>
<td>-2.81</td>
<td>.01</td>
</tr>
</tbody>
</table>

*(table continues)*
Table 10. (continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>$B$</th>
<th>SEB</th>
<th>Beta</th>
<th>$t$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>51.49</td>
<td>1.73</td>
<td>29.83</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>$Z'_1$</td>
<td>-3.89</td>
<td>2.98</td>
<td>-.08</td>
<td>-1.30</td>
<td>.19</td>
</tr>
<tr>
<td>$Z'_2$</td>
<td>1.76</td>
<td>1.66</td>
<td>.04</td>
<td>1.06</td>
<td>.29</td>
</tr>
<tr>
<td>$Z'_3$</td>
<td>-1.68</td>
<td>1.45</td>
<td>-.04</td>
<td>-1.16</td>
<td>.29</td>
</tr>
<tr>
<td>$Z_1$</td>
<td>-.77</td>
<td>2.10</td>
<td>-.01</td>
<td>-.37</td>
<td>.72</td>
</tr>
<tr>
<td>$Z_2$</td>
<td>1.96</td>
<td>2.15</td>
<td>.03</td>
<td>.91</td>
<td>.36</td>
</tr>
<tr>
<td>$Z_3$</td>
<td>6.33</td>
<td>1.84</td>
<td>.14</td>
<td>3.45</td>
<td>.001</td>
</tr>
<tr>
<td>$Z_4$</td>
<td>-.98</td>
<td>1.91</td>
<td>-.02</td>
<td>-.51</td>
<td>.61</td>
</tr>
<tr>
<td>$Z_5$</td>
<td>-.91</td>
<td>2.23</td>
<td>-.01</td>
<td>-.41</td>
<td>.68</td>
</tr>
<tr>
<td>$Z_6$</td>
<td>-8.10</td>
<td>2.88</td>
<td>-.08</td>
<td>-2.81</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>$Z''$</td>
<td>4.18</td>
<td>2.58</td>
<td>.10</td>
<td>1.62</td>
<td>.11</td>
</tr>
</tbody>
</table>

Note. Adjusted $R^2 = .007$ for Step 1; Δ Adjusted $R^2 = .030$ for Step 2; Δ Adjusted $R^2 = .002$ for Step 3 (ps = .02, < .001, .11, respectively).

Research Question 3

To what extent are differences between schools related to math achievement?

Research Question 3 utilized the second block (Step 2) of variables in the Table 10 hierarchical regression. In this step, differences between schools ($Z_1$ - $Z_6$) were added to the model to determine their impact on math achievement, net of year of implementation. The analysis demonstrated a significant ANOVA for Step 2, $F(9, 1434) = 6.07, p < .001$. From Step 1, the incremental increase of .03 in the Adjusted $R^2$ to .037 meets the Cohen et al. (2003) criterion of .02 minimal addition for large $N$. However, this is still a very small
The results varied from school to school demonstrating school differences. The $Z'_2$ contrast for Year 2 of implementation remains significant. The school contrasts for Schools F ($Z_3$) and I ($Z_6$) compared to the referent category, School C, were also significant. The largest standardized beta in Step 2 is .12 for $Z_3$, indicating a change of .12 standard deviation units on third grade NCE math scores for a one standard deviation change on the $Z_3$ contrast.

**Research Question 4**

To what extent is the beginning year of the program related to math achievement?

For Research Question 4, the relevant portion of the Table 10 hierarchical regression is Step 3 (Beginning Year of Implementation) which examines whether the Beginning Year (BGYR) in which the schools began the America’s Choice program has an impact on the mathematics achievement scores, after the effects of year of implementation and schools have been controlled. The schools began the program at different times representing 2002 for Schools C, G-I, and 2003 for Schools D-F. The ANOVA for Step 3 is significant, $F(10, 1433) = 5.73$, $p < .001$. However, $Z''$ is non-significant and the addition of beginning year produces only a .002 change in the Adjusted $R^2$ (.032 for Step 3), less than the .02 minimal increment for the criterion of Cohen et al. (2003). For Step 3, none of the three year-of-implementation $Z'$ contrasts were significant while the $Z_3$ and $Z_6$ contrasts for schools remain significant. For Step 3 the largest standardized beta is .14 for $Z_3$. Thus in the full model, the entire effect comes from the difference for two schools (F and I) compared to the referent category, School C. (The first listed school was the variable against which the other schools would be compared.)

**Research Question 5**

What is the effect of controlling for demographic factors for Research Questions 2-4?
Research Question 5 utilized hierarchical regression to examine the additional effect of controlling for demographic factors during the analyses for Research Questions 2-4. Four steps were included to examine the different blocks of variables sequentially with demographic factors (size and percent free/reduced lunch) entered in Step 1, implementation factors dummy coded as \( Z' \) contrasts added in Step 2, dummy coded \( Z \) contrasts for differences across schools added in Step 3, and beginning year of implementation \( Z'' \) contrast entered last in Step 4. The simultaneous entry within each step calculated the contribution of each independent variable, including the raw and standardized regression coefficients as well as the \( t \) scores and statistical significance.

Table 11 presents the results of the hierarchical regression model for all four steps. For Step 1, the demographic variables size (SIZE) and percent free/reduced lunch (%F/RL) were entered into the model to determine their effect on math achievement scores. The model was non-significant \( F(2, 1029) = 0.065, \ p = .94 \). The Adjusted \( R^2 \) of .00 indicates no effect on mathematics scores for the two demographic factors. Of note is the missing data for the demographic variables (\( N = 1031 \) for Table 11 vs. \( N = 1443 \) for the implementation model variables in Table 10). Furthermore, all schools had high percentages of students participating in the free/reduced lunch program (see Table 2, Chapter III), but for third grade only, the mean percentages for free/reduced lunch were considerably lower, ranging from \( M = 12\% \) for School F to \( M = 79\% \) for School I (see Table 6, Chapter IV). Also, as noted in Table 6, for the year 2004, School I (\( n = 12 \)) did not include free and reduced lunch data in their report to the state. However, School I was the smallest so that the impact on the regression for these 12 cases should be minimal. Unknown, however, is the extent that the almost 300 cases of missing data for demographic variables affected the non-significant finding.
For Step 2 (Table 11), year of implementation (Year 1, Year 2, and Continuation Year) was added to the equation to determine the effect after controlling for demographic factors. The ANOVA for this model was non-significant $F(5, 1026) = 1.41, p = .22$. The Adjusted $R^2$ of .002 is still less than 1% after year of implementation was added.

For Step 3 of Table 11, differences between schools were added to the equation to consider their effect on NCE math scores when demographic factors were controlled. The results for the ANOVA model, $F(11, 1020) = 2.01, p = .03$, were significant. The effect size (1.1%) is still close to zero, consistent with a change in $R^2$ of less than two percent having no substantive meaning, per Cohen et al. (2003). With the addition of the school contrasts, Year 2 changed from significant to non-significant, but SIZE becomes significant and has the largest standardized beta at -.30. School contrasts $Z_3$, $Z_5$, and $Z_6$ demonstrated significance (compared to referent, School C). The largest standardized beta for Schools was $Z_3$ at .24 indicating a change of .24 standard deviation units on NCE mathematics scores.

In Step 4 of Table 11, beginning year of implementation was added to the equation in addition to demographic factors, year of implementation, and schools. Beginning year of implementation was 2003 for schools D-F and 2002 for schools C, G-I. The results for the Step 4 model, $F(12, 1019) = 1.91, p = .03$ were significant, but the Adjusted $R^2$ remains at 1% (actually a slight decrease of .001 from Step 3 to Step 4). The beginning year factor was non-significant. For Step 4, SIZE had the largest standardized Beta at -.25, with school contrasts (compared to referent School C) for $Z_3$, School F, and $Z_6$, School I, remaining significant; the $Z_5$ contrast changed from significant in Step 3 to non-significant. However, given the overall effect size, these levels of significance have no practical impact on math achievement. For Step 4, SIZE had the largest standardized
Beta at -.25, interpreted as a change of -.25 standard deviation units on the math scores associated with a one standard deviation change in the independent variable.

Table 11

*Hierarchical Regression of Mathematics NCE Scores on Demographic Variables and Three Blocks of America’s Choice Implementation Model*

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
<th>Beta</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>50.72</td>
<td>2.41</td>
<td>21.05</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>SIZE</td>
<td>.01</td>
<td>.03</td>
<td>.01</td>
<td>.31</td>
<td>.76</td>
</tr>
<tr>
<td>%F/RL a</td>
<td>-.35</td>
<td>1.89</td>
<td>-.01</td>
<td>-.19</td>
<td>.85</td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>49.24</td>
<td>2.73</td>
<td>18.05</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>SIZE</td>
<td>.02</td>
<td>.03</td>
<td>.02</td>
<td>.65</td>
<td>.52</td>
</tr>
<tr>
<td>%F/RL a</td>
<td>-.62</td>
<td>1.90</td>
<td>-.01</td>
<td>-.33</td>
<td>.74</td>
</tr>
<tr>
<td>$Z'_1$</td>
<td>.47</td>
<td>1.72</td>
<td>.01</td>
<td>.27</td>
<td>.79</td>
</tr>
<tr>
<td>$Z'_2$</td>
<td>4.01</td>
<td>1.70</td>
<td>.09</td>
<td>2.34</td>
<td>.02</td>
</tr>
<tr>
<td>$Z'_3$</td>
<td>.35</td>
<td>1.70</td>
<td>.01</td>
<td>.20</td>
<td>.84</td>
</tr>
</tbody>
</table>

*(table continues)*
Table 11. (continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>$B$</th>
<th>$SE\ B$</th>
<th>Beta</th>
<th>$t$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>65.33</td>
<td>5.54</td>
<td></td>
<td>11.80</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>SIZE</td>
<td>-.28</td>
<td>.10</td>
<td>-.30</td>
<td>-2.74</td>
<td>.01</td>
</tr>
<tr>
<td>%F/RL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-.33</td>
<td>1.90</td>
<td>-.01</td>
<td>-.17</td>
<td>.86</td>
</tr>
<tr>
<td>$Z'_1$</td>
<td>-1.28</td>
<td>1.82</td>
<td>-.03</td>
<td>-.70</td>
<td>.48</td>
</tr>
<tr>
<td>$Z'_2$</td>
<td>2.59</td>
<td>1.76</td>
<td>.06</td>
<td>1.47</td>
<td>.14</td>
</tr>
<tr>
<td>$Z'_3$</td>
<td>-1.33</td>
<td>1.76</td>
<td>-.03</td>
<td>-.75</td>
<td>.45</td>
</tr>
<tr>
<td>$Z_1$</td>
<td>.17</td>
<td>2.11</td>
<td>&lt;.001</td>
<td>.08</td>
<td>.94</td>
</tr>
<tr>
<td>$Z_2$</td>
<td>-1.77</td>
<td>2.24</td>
<td>-.03</td>
<td>-.79</td>
<td>.43</td>
</tr>
<tr>
<td>$Z_3$</td>
<td>16.20</td>
<td>6.61</td>
<td>.24</td>
<td>2.45</td>
<td>.01</td>
</tr>
<tr>
<td>$Z_4$</td>
<td>4.69</td>
<td>2.70</td>
<td>.10</td>
<td>1.74</td>
<td>.08</td>
</tr>
<tr>
<td>$Z_5$</td>
<td>-4.87</td>
<td>2.36</td>
<td>-.08</td>
<td>-2.06</td>
<td>.04</td>
</tr>
<tr>
<td>$Z_6$</td>
<td>-15.58</td>
<td>4.23</td>
<td>-.17</td>
<td>-3.68</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

(table continues)
Table 11. (continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
<th>Beta</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>62.52</td>
<td>6.33</td>
<td>9.88</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>SIZE</td>
<td>-.23</td>
<td>.12</td>
<td>-.25</td>
<td>-2.00</td>
<td>.05</td>
</tr>
<tr>
<td>%F/RL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-.29</td>
<td>1.90</td>
<td>-.01</td>
<td>-1.15</td>
<td>.25</td>
</tr>
<tr>
<td>Z&lt;sub&gt;1&lt;/sub&gt;</td>
<td>-3.93</td>
<td>3.40</td>
<td>-.08</td>
<td>-1.15</td>
<td>.25</td>
</tr>
<tr>
<td>Z&lt;sub&gt;2&lt;/sub&gt;</td>
<td>1.55</td>
<td>2.10</td>
<td>.03</td>
<td>.74</td>
<td>.46</td>
</tr>
<tr>
<td>Z&lt;sub&gt;3&lt;/sub&gt;</td>
<td>-1.16</td>
<td>1.77</td>
<td>-.03</td>
<td>-.66</td>
<td>.51</td>
</tr>
<tr>
<td>Z&lt;sub&gt;1&lt;/sub&gt;</td>
<td>.62</td>
<td>2.16</td>
<td>.01</td>
<td>.29</td>
<td>.78</td>
</tr>
<tr>
<td>Z&lt;sub&gt;2&lt;/sub&gt;</td>
<td>-.91</td>
<td>2.43</td>
<td>-.02</td>
<td>-.38</td>
<td>.71</td>
</tr>
<tr>
<td>Z&lt;sub&gt;3&lt;/sub&gt;</td>
<td>13.79</td>
<td>7.11</td>
<td>.20</td>
<td>1.94</td>
<td>.05</td>
</tr>
<tr>
<td>Z&lt;sub&gt;4&lt;/sub&gt;</td>
<td>3.83</td>
<td>2.86</td>
<td>.08</td>
<td>1.34</td>
<td>.18</td>
</tr>
<tr>
<td>Z&lt;sub&gt;5&lt;/sub&gt;</td>
<td>-3.90</td>
<td>2.59</td>
<td>-.07</td>
<td>-1.51</td>
<td>.13</td>
</tr>
<tr>
<td>Z&lt;sub&gt;6&lt;/sub&gt;</td>
<td>-14.31</td>
<td>4.45</td>
<td>-.16</td>
<td>-3.22</td>
<td>.001</td>
</tr>
<tr>
<td>Z&lt;sup&gt;*&lt;/sup&gt;</td>
<td>2.97</td>
<td>3.23</td>
<td>.07</td>
<td>.92</td>
<td>.36</td>
</tr>
</tbody>
</table>

Note. Adjusted $R^2 = -.00$ for Step 1; Adjusted $R^2 = .007$ for Step 2; Adjusted $R^2 = .014$ for Step 3; Adjusted $R^2 = .001$ for Step 4 ($p = .937, .075, .021, .358$ respectively)

<sup>a</sup>School I did not report Free/reduced data for 2004; the impact was minimal because School I is the smallest (see Table 5--$n = 12$ for missing Cl data).

Therefore, the effect of controlling for demographic factors (SIZE and %F/RL) had little or no impact on math achievement when year of implementation, differences
between schools, and beginning year were added to the model. Schools F and I (the largest and smallest by population) were significant.

Research Question 6

For the three years of the America’s Choice mathematics program (Year 1, Year 2, Continuation 1), how does the progress of the America’s Choice schools compare to that statewide for the same period?

Research Question 6 examined comparative change for the seven America’s Choice schools compared to the state for Year 1 and Year 2 of Implementation, plus the Continuation Year. For schools C, G, H, and I the Prior Year (P1) was 2001. Prior Year was 2002 for Schools D-F. Research Question 6 was addressed in two ways. First, one sample t tests were conducted to compare mathematics achievement for the America’s Choice schools to the statewide scores for Year 1, Year 2, and the Continuation Year. (The baseline Prior Year was compared in RQ1.)

For RQ6, only the collective progress of the America’s Choice schools is calculated (separately for Schools C, G-I and for Schools D-F because of staggered start dates for this project). Individual school effects were examined in Research Question 3, above. Table 12 shows the results of the t tests conducted to compare the third grade NCE mean math scores for the America’s Choice schools and the state for Year 1 and Year 2 of implementation plus the Continuation Year.

It seems that Schools D-F were not significantly different in for Year 1 and Year 2 of implementation, but were significantly below the state average for the Continuation Year. Schools C, G-I were significantly different from the state math mean for Year 1 of implementation but not for Year 2 and were below the state for the Continuation Year but not significantly so.
Table 12

One Sample t Tests Comparing Mean NCE Third Grade Mathematics Scores

for America’s Choice Versus the State for Year 1, Year 2, and Continuation Year

<table>
<thead>
<tr>
<th>Schools</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>t</th>
<th>p</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C, G-I(^a)</td>
<td>149</td>
<td>50.3</td>
<td>20.43</td>
<td>-2.89</td>
<td>&lt;.01</td>
<td>-8.15</td>
<td>-1.53</td>
</tr>
<tr>
<td>D-F(^b)</td>
<td>209</td>
<td>55.4</td>
<td>18.46</td>
<td>-1.29</td>
<td>.20</td>
<td>-4.16</td>
<td>.88</td>
</tr>
<tr>
<td>Year 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C, G-I(^c)</td>
<td>150</td>
<td>54.0</td>
<td>22.70</td>
<td>-1.62</td>
<td>.11</td>
<td>-12.23</td>
<td>1.33</td>
</tr>
<tr>
<td>D-F(^d)</td>
<td>201</td>
<td>56.7</td>
<td>20.42</td>
<td>-1.27</td>
<td>.21</td>
<td>-4.67</td>
<td>1.01</td>
</tr>
<tr>
<td>Continuation Year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C, G-I(^e)</td>
<td>160</td>
<td>50.6</td>
<td>20.77</td>
<td>-4.83</td>
<td>&lt;.01</td>
<td>-11.17</td>
<td>-4.68</td>
</tr>
<tr>
<td>D-F(^f)</td>
<td>193</td>
<td>51.2</td>
<td>19.0</td>
<td>-5.97</td>
<td>&lt;.01</td>
<td>-10.85</td>
<td>-5.46</td>
</tr>
</tbody>
</table>

\(^a\)For Schools C, G-I, Y1 of four years was 2002 (µ = 55.1).

\(^b\)For Schools D-F, Y1 of four years was 2003 (µ = 57).

\(^c\)For Schools C, G-I, Y2 of four years was 2003 (µ = 57).

\(^d\)For Schools D-F, Y2 of four years was 2004 (µ = 58.5).

\(^e\)For Schools C, G-I, Continuation Year was 2004 (µ = 58.5).

\(^f\)For Schools D-F, Continuation Year was 2005 (µ = 59.3).

In the second approach to analyzing Research Question 6, percentage change scores were calculated and compared for America’s Choice versus the state. That analysis required multiple steps. Step 1 computed the separate means for both America’s Choice and statewide consistent with the staggered start dates for the project schools. The NCE
Grade 3 Mathematics scores for each of the seven America’s Choice schools were reported previously (see Table 7).

Table 13 shows the figures used in the percentage change calculations, including the overall sample means for the four America’s Choice Schools that began in 2001 and the three that began in 2002, as well as the population means for the state for the two four-year periods. Schools C, G-I showed an increase from Prior Year (P1) to Year 2 (Y2) of implementation, and then regressed during the Continuation Year (C1) to the level for Year 1. The mean scores for Schools D-F increased slightly from Prior Year (P1) through Year 2 (Y2) of implementation, then demonstrated a decrease from Year 2 to the Continuation Year (C1) to below the Prior Year level. Both groups of schools were performing below the state throughout the intervention. The statewide results for 2001-2004 and 2002-2005 showed a steady increase for all four years.
In Step 2, the author calculated percentage change scores (%Change) based on CTBS/5 third grade NCE math scores for the America’s Choice schools and the state. Individual student change scores cannot be computed for these data because Kentucky’s accountability model is based on comparison across successive cohorts, not a longitudinal growth model. Thus, figures in Table 13 represent the collective means for the groupings of schools indicated. Accordingly the percentage change scores in Table 14 are based on comparison of these groupwide means. The percentage change scores were calculated for the mean math scores using the following formula: (target year minus Prior Year) divided by Prior Year. The change scores were calculated for Year 1 minus Prior Year, Year 2 minus Prior Year, and the Continuation Year minus Prior Year.

Table 14 shows the unweighted percentage change scores (%Change) for the two groups of America’s Choice schools and the state. (All calculations in Table 14 are for the

Table 13

Mean NCE Math Scores Utilized for Cohort Percentage Change Calculations for America’s Choice and Statewide

<table>
<thead>
<tr>
<th>Unit</th>
<th>P1</th>
<th>Y1</th>
<th>Y2</th>
<th>C1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
<td>N</td>
<td>M</td>
</tr>
<tr>
<td>C, G-I\textsuperscript{a}</td>
<td>186</td>
<td>48.2</td>
<td>149</td>
<td>50.3</td>
</tr>
<tr>
<td>D-F\textsuperscript{b}</td>
<td>190</td>
<td>54.9</td>
<td>209</td>
<td>55.4</td>
</tr>
<tr>
<td>Statewide\textsuperscript{a}</td>
<td>49,800</td>
<td>54</td>
<td>48,764</td>
<td>55.1</td>
</tr>
<tr>
<td>Statewide\textsuperscript{b}</td>
<td>48,764</td>
<td>55.1</td>
<td>48,120</td>
<td>57</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Schools C, G-I and statewide third grade for 2001-2004.

\textsuperscript{b}Schools D-F and statewide third grade for 2002-2005.
target year compared back to the Prior Year of implementation. The change scores for
Year 1 to Year 2 of the America's Choice mathematics component and for Year 2 to the
Continuation Year were not calculated.) For Schools C, G, H, and I the calculations
showed an increase from their Prior Year to Year 1 and a much bigger gain for Year 2. For
the Continuation Year there is an increase compared to the Prior Year but the considerably
smaller percentage change reflects the decline in actual achievement from Year 2 to
Continuation Year. Schools D-F, when compared to the Prior Year, showed a slight gain
for Year 1 with a larger increase for Year 2 and a large decline from Year 2 to the
Continuation Year, as reflected in Table 13. For the state, the percentage change
(compared to the Prior Year) increased for each subsequent year for both the 2002 and
2003 start dates.

Table 14

Percentage Change Scores across Staggered Start Dates for Year 1, Year 2, and
Continuation Year for America's Choice Schools and Statewide

<table>
<thead>
<tr>
<th>Unit</th>
<th>Year 1 2002</th>
<th>Year 1 2003</th>
<th>Year 2 2003</th>
<th>Year 2 2004</th>
<th>Continuation Year 2004</th>
<th>Continuation Year 2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC C, G-I</td>
<td>4.36</td>
<td>12.03</td>
<td></td>
<td></td>
<td>4.98</td>
<td></td>
</tr>
<tr>
<td>AC D-F</td>
<td>.91</td>
<td>3.28</td>
<td></td>
<td></td>
<td>-6.74</td>
<td></td>
</tr>
<tr>
<td>State</td>
<td>2.03</td>
<td>5.55</td>
<td></td>
<td></td>
<td>8.33</td>
<td></td>
</tr>
<tr>
<td>State</td>
<td>3.44</td>
<td>6.17</td>
<td></td>
<td></td>
<td>7.62</td>
<td></td>
</tr>
</tbody>
</table>

Note. AC C, G-I = America's Choice Schools C, G, H, and I; AC D-F = America's Choice Schools D, E, and F.

aPrior Year is 2001; %Change is for target years 2002-2004.
bPrior Year is 2002; %Change is for target years 2003-2005.
Next, in Step 3 *weighted* percentage change (%Change) scores were calculated for the seven America's Choice Schools and statewide for Year 1, for Year 2, and for Continuation Year (based on Step 2 results). Weighted %Change scores were calculated because the question addressed the overall impact of America's Choice. This required combining the two cohorts of schools (based on start dates) to be examined altogether.

The calculation used the following formula for America's Choice: \( \frac{(%\text{Change across Schools C, G-I})(N \text{ for Schools C, G-I}) + (%\text{Change across Schools D-F})(N \text{ for Schools D-F})}{(N \text{ for Schools C, G-I}) + (N \text{ for Schools D-F})} \). A similar formula was used in the calculation of the state *weighted* percentage change (%Change) scores based on the statewide populations for Prior Year 2001 and for Prior Year 2002, respectively. (These %Change score values for the seven schools, as compared to the state for Year 1, Year 2, and Continuation Year, were taken from Table 14). Table 15 presents these weighted percentage change scores for both America's Choice Schools and statewide.

Table 15

*Weighted Percentage Change Scores for Seven America's Choice Schools and for Statewide for Year 1, Year 2, and Continuation Y*

<table>
<thead>
<tr>
<th>Unit</th>
<th>Y1(^a)</th>
<th>Y2(^b)</th>
<th>C1(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>2.35</td>
<td>7.02</td>
<td>-1.44</td>
</tr>
<tr>
<td>State</td>
<td>2.73</td>
<td>5.86</td>
<td>7.98</td>
</tr>
</tbody>
</table>

*Note. AC = Seven America's Choice Schools.*


The trend in Table 15 is notable: for the seven America’s Choice schools, the intervention result is slight improvement for Y1 (about the same as the State) followed by considerable improvement for Y2 (greater than the State), but then a precipitous drop for the Continuation Year, ending at a level below that of the Prior Year baseline scores (compared to a continuing increase for the State). Figure 4, graphing the rate of progress over the two years of implementation and the continuation year, presents a stark visual of the State’s continued progress versus the decline in the sustainability year for the America’s Choice sample. The line for discrepancy shows this another way (State minus America’s Choice); after the Continuation Year, the State rate of change (%Change scores) is about 9.5% ahead of America’s Choice with the entire advantage accruing during the year after active implementation of the program.

Figure 4. Comparing weighted %Change Scores for America’s Choice versus the State.

Summary
The purpose of this study was to investigate the impact of implementing the mathematics component of America’s Choice comprehensive reform. To explore this, the study analyzed Comprehensive Test of Basic Skills (CTBS/5) Normal Curve Equivalent (NCE) grade three mathematics data, collected from the Kentucky Department of Education. The central question guiding this study was: What are the effects of implementing America’s Choice comprehensive school reform design in mathematics in seven elementary schools in Kentucky. This chapter presented the findings of the analyses for each of the six specific Research Questions that guided the study.

Data collected over four cohorts represented the year prior to implementation (P1), two years of supported implementation (Y1 and Y2), and the Continuation Year (C1) with no direct support. The analysis included descriptive statistics for the control demographic variables (size and percent free/reduced lunch); the implementation model variables (year of implementation, schools, and beginning year of implementation) were dummy coded so no descriptives were reported. The dependent variable was the Comprehensive Test of Basic Skills (CTBS/5) Normal Curve Equivalent (NCE) grade three mathematics scores.

The study was confined to Kentucky elementary schools implementing the mathematics component of the America’s Choice comprehensive reform model. While nine schools selected America’s Choice, only seven implemented the mathematics component for the years included in the study.

Research Question 1 utilized one sample t tests to compare Comprehensive Test of Basic Skills (CTBS/5) Normal Curve Equivalent (NCE) mathematics mean scores from the America’s Choice schools to the state for the year before the project was begun (Prior Year) of the study.

Two separate hierarchical regressions were designed for this study. The first
addressed Research Questions 2-4 examining the effects of three dimensions of implementation; change over years (RQ2), differences between schools (RQ3), and different beginning years (RQ4). The predictor variables were entered in three steps based on nominal coding of Z Contrasts as explained in Chapter III with simultaneous entry within each step: Year of Implementation (Year 1, Year 2, and Continuation Year entered as $Z_1'$, $Z_2'$, and $Z_3'$, respectively, with Prior Year as the referent); Schools D-I (entered as $Z_7$-$Z_6$) with School C as the referent; and Beginning Year of Implementation entered as $Z''$.

A separate hierarchical regression addressed Research Question 5 which examined the effect of controlling for demographic factors for RQs 2-4. The variables were entered in four steps. In Step 1, demographic variables (size and percent free/reduced lunch) were entered into the model. Step 2 added the implementation variables: two years of implementation ($Y_1$ and $Y_2$), and Continuation Year ($C_1$) with Prior Year as the referent category. Step 3 included the addition of differences across schools. The Beginning Year of Implementation ($BGYR$) was added in Step 4. The variables were entered simultaneously within each of the models, utilizing nominal coding of Z contrasts from Chapter III. The dependent variable was the Comprehensive Test of Basic Skills (CTBS/5) Normal Curve Equivalent (NCE) mathematics scores for third grade exiting primary students.

For Research Question 6, the analysis required two approaches. First, one sample $t$ tests compared the mean NCE mathematics scores for the schools versus the state for Year 1, Year 2, and Continuation Year. The second approach computed percentage change scores over the course of the implementation and required several steps. Step 1 included the calculation of math means for the America's Choice schools and statewide consistent with the staggered start dates for the two groups of schools (see Table 13). Step 2 required
the calculation of percentage change scores (%Change) for the America’s Choice schools and the state for Year 1, Year 2, and Continuation Year using the following formula: (target year minus the Prior Year) divided by the Prior Year (see Table 14). Step 3 included the calculation of weighted percentage change (%Change) scores for the seven America’s Choice Schools and statewide for Year 1, for Year 2, and Continuation Year using the formula: (%Change across Schools C, G-I)(N for Schools C, G-I) + (%Change across Schools D-F)(N for Schools D-F) / (N for Schools C, G-I) + (N for Schools D-F). A similar formula was used for the state calculations for Prior Year 2001 and for Prior Year 2002, respectively (see Table 15). A graph was then computed to visualize the rate of progress on percentage change scores for America’s Choice schools versus the percentage change scores for the state over the same period.

The overall results of the analysis for this study were somewhat disappointing with regards to the guiding question, What are the effects of implementing America’s Choice comprehensive school reform design in mathematics in seven elementary schools in Kentucky. Based on Table 15, for Year 1 of Implementation (Y1) the America’s Choice Schools increased at almost the same rate as the state with regard to mathematics achievement. Year 2 of Implementation (Y2) showed considerable progress in math scores for the project schools, actually increasing faster than the state as a whole. However, despite the gains the schools remained below the state mean. During the Continuation Year (C1), the America’s Choice schools demonstrated a negative change (to a level below the beginning of the project in the Prior Year) while the state continued to progress. These results agree with the one sample t tests reported in Table 12, i.e., the America’s Choice schools are below the state, bordering on significance for Year 1 and Year 2, but they are significantly lower for the Continuation Year.
CHAPTER V
DISCUSSION AND CONCLUSION

Introduction

This chapter is divided into four major sections: the study in brief; discussion of the findings for each research question including analysis and implications; recommendations regarding policy, practice, and future research; and conclusions.

The Study in Brief

The United States government has been concerned with the ability of America’s schools to provide high quality educational experiences to prepare all students to be successful after graduation, especially from the 1950s through today. To address this issue various reform initiatives were established to provide guidance for schools. From the Elementary and Secondary Education Act (ESEA) of 1965 to the more recent version, No Child Left Behind Act (NCLB) of 2001, school reform initiatives led to changes in the educational systems across America.

Kentucky redesigned its educational system with the development of the Kentucky Education Reform Act (KERA) in 1990. The goal of KERA was to provide high quality experiences within the national movement toward standards-based educational reform that continues under the No Child Left Behind Act. As part of the KERA initiative, the Standards and Indicators for School Improvement: Kentucky’s Model for Whole School Improvement (SISI) document (KDE, 2004a), along with the Scholastic Audit process helped schools measure their ability to increase student
achievement. Schools identified as needing assistance used some of their Title I funds to implement various reform models.

Comprehensive reform models focus on schoolwide improvements that address all aspects of school operations in an effort to raise student achievement, particularly for low achieving students. Title I funds are available to schools to implement a comprehensive model to address needs identified by the scholastic audit process developed by the Kentucky Department of Education (KDE). Options available for Kentucky schools to select from included: Accelerated Schools Project, America's Choice, ATLAS Communities, Early Intervention in Reading, First Steps, Modern Red Schoolhouse, School Development Program, and Success for All (St. John et al., 2000).

The National Center on Education and the Economy (NCEE) developed America's Choice comprehensive reform as a grade K through 12 model to address academic needs. This design provides a rigorous, standards-based curriculum and safety nets for students to help them achieve academic success. Five basic components of America's Choice include: high performance leadership and management; capacity building through aligned instructional systems; professional learning communities; improved student achievement for at-risk learners through standards and assessments; and parent and community involvement (Allen et al., 2003). Nine Kentucky elementary schools adopted America’s Choice comprehensive reform.

This research provided additional information about the effects of implementation of comprehensive reform models in elementary schools. Specifically, this built upon the work of VanMeter (2005) who studied the impact of implementing the reading component of America’s Choice comprehensive reform in nine Kentucky elementary schools. The central research question for this study was: What are the effects of
implementing America’s Choice comprehensive school reform design in mathematics in seven elementary schools in Kentucky? More specifically, the researcher looked at the following Research Questions:

For the seven America’s Choice schools for third grade mathematics:

1. To what degree do the America’s Choice schools differ from the statewide mean for math achievement at the beginning of the program (Prior Year)?

2. To what extent is the year of implementation related to math achievement?

3. To what extent are differences between schools related to math achievement?

4. To what extent is the beginning year of the program related to math achievement?

5. What is the effect of controlling for demographic factors for Research Questions 2-4?

6. For the three years of the America’s Choice mathematics program (Year 1, Year 2, Continuation 1), how does the progress of the America’s Choice schools compare to that statewide for the same period?

This quantitative study analyzed the Comprehensive Test of Basic Skills (CTBS/5) Normal Curve Equivalent (NCE) mathematics scores for third grade exiting primary students in the schools implementing America’s Choice comprehensive reform that had adequate data (seven of the nine). Due to the staggered start dates, data collected represented a five year period: year prior to implementation (P1), two years during implementation (Y1 and Y2), and the continuation year (C1) after supportive implementation ended. The schools began the program at different times so year prior to implementation was 2002 for schools C, G-I and 2003 for schools D-F. Scores from the year prior to implementation provided baseline scores for each school. The Commonwealth Accountability Testing System (CATS) utilized a cohort model with comparisons made across successive groups for each grade level tested. Accordingly, the
scores for this study do not constitute longitudinal data; the third graders for each of the relevant years represent a proxy for growth for each school during the implementation project.

After computing descriptive statistics for all variables, the analyses included two hierarchical regressions—a single regression model to investigate the effects of the implementation model for Research Questions 2-4 with the Step 1 block (change over years of implementation) representing RQ2; Step 2 (school differences) for RQ3; and Step 3 (beginning year of America’s Choice program) for RQ4. A separate hierarchical regression for Research Question 5 examined the effects of the implementation model (the three steps for the first hierarchical regression, now entered as Steps 2-4) while controlling for demographic factors which were entered in Step 1. Research Questions 1 and 6 used one sample \( t \) tests to compare progress of the America’s Choice schools to the state. (Each research question is listed in the discussion that follows). All analyses were conducted using the Statistical Package for Social Sciences (SPSS) computer program.

Independent variables included two types: control (demographics—size and percentage free/reduced lunch) and implementation model (year of implementation, differences between schools, and beginning year of implementation). The dependent variable was the Comprehensive Test of Basic Skills (CTB/5) Normal Curve Equivalent (NCE) mathematics scores for third grade exiting primary students. The sample consisted of seven Kentucky elementary schools implementing the mathematics component of America’s Choice comprehensive reform.

Discussion

This section addresses the findings of the study. Although results varied from school to school, the overall conclusion is that the America’s Choice Schools made some
gains during the two years of active implementation of the mathematics component. However, these gains were mostly lost during the year of continuation (first year after active implementation) and in some of the schools, the level of achievement fell below where they had started two years earlier.

In the discussion that follows each Research Question is treated separately to address briefly the findings followed by the analysis. For each Research Question, the findings are recapitulated. Then the implications of those results are analyzed.

*Descriptive Statistics*

In Kentucky, seven elementary schools implemented the mathematics component of America’s Choice and were included in this study. All seven were predominately white, low socio-economic schools in rural settings across the state. During the four years of this study, the third grade student populations for five of the schools (C, D, E, G, and H) ranged from 155 to 251. The two outliers were School I with 62 and School F with 420 students (see Table 5 in Chapter III).

The percentage of students participating in the federal free/reduced lunch program schoolwide ranged from a low 66% for School E to 98% for School I (see Table 2 in Chapter III). However, the reported percentages for participation for third grade were lower, ranging from 12% for School F to 79% for School I (Table 6, Chapter IV).

The Comprehensive Test of Basic Skills (CTBS/5) mathematics mean scores for the seven schools demonstrated the need for an intervention or assistance for math instruction. When the schools began implementing the math component of America’s Choice, all except School F were performing below the state for that year. During the two years of active implementation the schools experienced a mixed response to the intervention methods with most schools experiencing gains the first two years of
implementation. That progress was lost during the following year as these schools slipped back to where they had started or below. Because the state was making steady progress during this period, the schools began the America’s Choice below the state mean and fell even further behind during the program, primarily because of their dismal performance in the first year following active participation.

Analysis

All of these schools were low achieving and at a level below the state, Tier 3 Assistance. The one school that was an anomaly in this set of seven was School F which was also different in the fact that they chose to implement America’s Choice before receiving Blue Ribbon status for their efforts in closing the achievement gap. (This occurred during the first year of implementation of the math component.) They continued implementing America’s Choice for two more years (i.e., Year 2 and Continuation Year), during which time their scores fluctuated above and below the state mean which demonstrated their need for this program.

The findings of this research on mathematics achievement paralleled the results of the VanMeter (2005) study on the impact of reading achievement from the America’s Choice program in the same sample of Kentucky elementary schools. The impact of the reading program was also varied during the years of implementation and the continuation year. Other researchers (e.g., Borman et al., 2003; Erlichson, 2005; May & Supovitz, 2006) also noted that demographic factors may impact the results of the intervention implementation of comprehensive reform models in elementary schools.

Research Question 1

To what degree do the America’s Choice schools differ from the statewide mean for math achievement at the beginning of the program (Prior Year)?
Research Question 1 utilized one sample *t* tests to compare the Comprehensive Test of Basic Skills (CTBS/5) Normal Curve Equivalent (NCE) mathematics mean for each of the seven America's Choice schools individually and as cohorts versus the state. The analyses were conducted for the Prior Year to implementation of the math component of America's Choice. Based on the different beginning years for the program Schools C, G, H, and I were compared to the state for the year 2001 while Schools D-F were compared to 2002. Table 9 presented in Chapter IV shows the findings.

Six of the seven America's Choice schools were performing below the state mathematics mean for the year prior to implementation; the difference was significant, however, only for School G. For Prior Year, only School F performed above the state, significantly so. Schools C, G-I collectively were achieving significantly below the state, but schools D-F were not significantly distinct from the state mean. The latter finding is consistent with the fact that Schools D and E were below the mean while School F was the only one of the seven that was above the state.

*Analysis*

That all but one of the America's Choice schools were performing below the state average in mathematics achievement for the year prior to implementation of the math component of America's Choice was to be expected; all of these schools chose to participate in an intervention designed to help struggling schools improve their achievement so that they could meet the long term goal of "proficient" by the year 2014, as required under the Kentucky Education Reform Act of 1990 (KERA). Yet at the beginning of the Title I comprehensive reform initiative, only one of the seven was significantly lower than other schools statewide.

That conclusion, however, needs to be seen in context because Kentucky's math
achievement statewide was at an unacceptably low level. Being “lower but not significantly so” implies that change was still necessary, consistent with statewide efforts to precipitate the math reforms. Throughout the Commonwealth of Kentucky, mathematics development was a concern for educators due to the lagging performance of students when compared to other nations.

The implications of No Child Left Behind (NCLB) and Kentucky Reform initiatives could have affected the teachers’ focus on raising student achievement scores to be in compliance with state projected improvement goals. Schools in Tier 3 Assistance, based on the results of Scholastic Audits, examined various comprehensive reform models in order to select the appropriate model that best met their academic needs. For schools struggling with mathematics instruction, America’s Choice reform appeared to be a good selection because of the math workshop format that is structured to provide opportunities for students to engage actively in problem solving and discussions to develop a deeper understanding about math concepts.

Research Question 2

To what extent is the year of implementation related to math achievement?

A single hierarchical regression consisting of three sequential blocks for entering variables (simultaneous within each step) was conducted to address Research Questions 2-4. These questions examined the impact of the three dimensions of implementation on the Comprehensive Test of Basic Skills (CTBS/5) Normal Curve Equivalent (NCE) mathematics scores for third grade students in the Kentucky schools implementing America’s Choice reform. Step 1 addressed change over years of implementation (RQ2), Step 2 added differences between schools (RQ3), and Step 3 different beginning years (RQ4). Table 10 in Chapter IV provides the results of these analyses.
Research Question 2 addressed the second dimension of implementation, change over years of implementation. The researcher used Z' contrasts to compare Year 1, Year 2, and Continuation Year to the Prior Year to determine the impact of America’s Choice on math achievement in seven elementary schools. The findings from the overall model were significant for Step 1; however, the Adjusted $R^2$ of .01 indicates essentially no effect due to implementation year on mathematics achievement. For analyses with large $N$, the combined contrasts for each year were not close to the Cohen et al. (2003) criterion of .02 required for incremental change to be substantive.

*Analysis*

The beginning year of implementation for the math component occurred during the second year of program implementation. Reading and Writing were introduced during the first year of implementation of America’s Choice (which was the prior year for math). Supovitz and May (2004b) stated that average gains in student learning occurred when teachers implemented at least eight components of the program. With math implementation starting in the second year, most teachers were already familiar with the basic components and the elements of the workshop format of America’s Choice. For this intervention, however, the math component was apparently different such that gains in math did not really show up until Year 2 of math implementation (Year 3 of the overall project), albeit not enough for the year of implementation dummy contrasts to produce a significant effect. The minimal gains from Year 2, however, were lost during the Continuation Year (see Table 13, Chapter IV).

Two interpretations of this outcome are relevant. The first is simply a level of use explanation. Any intervention is only as good as the quality of its implementation, and there is considerable evidence from both the comprehensive school reform literature (cf.
McCaslin et al., 2008; Supovitz & May, 2004a; Wetherill & Applefield, 2005) and the work on implementation and change (see Supovitz & May, 2004b; Sterbinsky et al., 2005; Zhang et al., 2006) that what/how teachers do with a new program is what counts (extent of change in attitudes, beliefs, skills, instructional strategies, etc.). Based on research, it is known that it takes several years to realize gains from the implementation of any intervention (Beam & Faddis, 2001; Carlson, 2003; Erlichson, 2005).

The second explanation is an extension of the first, only based on evidence from this study. Table 14 gives percentage change for Year 1, Year 2, and Year 3, compared to the Prior Year, broken down for the two groups of schools. It is evident that the first group (Schools C, G, H, and I) made progress; Schools D-F, however, actually regressed. Because Research Question 1 examined the seven America’s Choice Schools as a whole, any separate effects would have almost certainly been washed out. Table 14, then, suggests that the likeliest explanation for the “no effects” of RQ1 is differences between schools that began implementation of America’s Choice in 2002 versus those that began in 2003. These data do not provide evidence on why this discrepancy occurred. Research Question 3 (next) addresses this issue of school to school differences directly.

Research Question 3

To what extent are differences between schools related to math achievement?

Research Question 3 utilized Step 2 of the hierarchical regression (see Table 10 in Chapter IV) to examine the second dimension of implementation: differences across schools. Z contrasts were utilized to compare the schools to determine the impact on math achievement. The results for the overall model in Step 2 were significant; the Adjusted $R^2$ of .03 is small but significant and exceeds the .02 minimum suggested by Cohen et al. (2003). The school contrasts for $Z_3$, and $Z_6$ were significant, but in opposite directions.
Compared to the referent, School C, School F improved and School I declined. This variation occurred within the context of a set of low performing schools, Tier 3 Assistance, the lowest rated in the state with respect to Kentucky’s expectation of continued progress. (School F was the exception; their opting into the program was a strategy for improvement but not required.)

**Analysis**

The study examined rural, predominately white (90% in School C to 100% in School I), low socio-economic elementary schools located across the Commonwealth of Kentucky. The schoolwide populations ranged from 154 students in School I to 642 in School F. All had varying percentages of students participating in the federal free/reduced lunch program (38% in School F to 98% in School I). The results may have looked different had a larger sample of elementary schools with more variation been included in the study. However, this study only focused on the Kentucky elementary schools that chose to adopt America’s Choice.

The fact that math achievement varied from school to school is not surprising. The literature on school reform and implementation models is replete with evidence that hoped for improvements are uneven and are related to the extent that faculty actually change what they are doing (Beam & Faddis, 2001; Holdzkom, 2002; Supovitz & May, 2004a; Supovitz & May, 2004b). Research findings by VanMeter (2005) also noted school to school differences when examining reading achievement in the nine Kentucky schools implementing America’s Choice (the seven included in this math study were part of VanMeter’s sample). VanMeter’s (2005) results showed Schools G and I were unstable, Schools F and H declined all years, while Schools C-E increased reading achievement scores. Thus these America’s Choice schools were also inconsistent in the degree of
progress (or regress) that each made for reading and math. Schools that did well in one of the two areas did not necessarily do well on the other. This implies that changes were occurring teacher by teacher but not necessarily for the entire school.

Furthermore, the differences between schools are compounded by the uneven results from year to year during and after the intervention (see analysis for RQ2, above). And to the extent that the schools saw slight improvements in math achievement during the years of supported implementation, they were unable to sustain this growth once support for implementation ended, as evidenced in the drop in scores during the continuation year, a trend across almost every school.

Whether this trend speaks to low levels of implementation, failure of the newer math methods to endure once active support was taken away, or other endemic problems related to the school learning climate or other factors is unknown. Very likely some combination of these issues was influential in why these schools were in the Tier 3 Assistance group for Kentucky (those identified as making the least progress toward the statewide goal of Proficient for all schools by 2014). Follow up qualitative case studies of these seven schools could be very instructive as to the specific dynamics within and across these America’s Choice sites.

Research findings from Supovitz et al. (2002) and Supovitz and May (2004a, 2004b) demonstrate that student gains begin slowly during the first year of implementation and gradually increase through the third year or continuation year. This trend was noted in the results for School C. School G saw an increase during the Year 1 and Year 2 of active mathematics implementation but declined during the continuation year (see Table 7). In fact, inspection of school to school progress across the three years of the program data reveals clear-cut inconsistency for both school to school and year to year. Part of the
explanation for this could be that with the exception of School F, all are relatively small, with the number of third graders ranging from 12 to 78. (School F had more than 100 in each year of implementation.) Cohort variation can be a very real phenomenon with such small numbers.

That being said, research consistently shows that instruction has a strong effect on achievement outcomes (Hiebert & Wearne, 1993; McKinney, 2007; Michael & Young, 2005; Murphy, 2004). The Kentucky educational directives focus on strengthening instructional practices to impact achievement outcomes. Teacher’s implementation of the various instructional practices is known to vary within schools and between schools. The current study of America’s Choice mathematics implementation supports the findings that variance occurs between schools when implementing changes (Beam & Faddis, 2001; Sterbinsky et al., 2003; VanMeter, 2005).

Research Question 4

To what extent is the beginning year of the program related to math achievement?

Research Question 4 utilized a Z" contrast in Step 3 of the hierarchical regression (see Table 10 in Chapter IV) to examine the third dimension of implementation: the effect of different beginning years on math achievement for the seven schools combined. For Schools C, G, H, and I the first year of implementation was 2002. For Schools D-F, the first year was 2003. The model demonstrated a significant effect. However, the Adjusted $R^2$ of .032 was an increase of only .002 from Step 2, less then the minimal criterion of .02 (Cohen et al., 2003).

Analysis

Since the implementation of Kentucky Education Reform Act in 1990 (KERA),
statewide achievement averages in public education have increased on the state assessment with the elementary level demonstrating the largest percentage increase. Third grade Comprehensive Test of Basic Skills (CTBS/5) results have increased from the 50th percentile in 1997 to the 65th percentile in 2004 (KDE, 2004). This statewide increase could have impacted the mathematics scores in these schools in addition to the changes resulting from implementation of America’s Choice.

Kentucky’s increase in achievement since 1990 has been well documented (cf. Eisner, 2001; Poggio, 2000; Rothstein, 2004). Further, this trend has been demonstrated to have a significant impact on Scholastic Audit data collected by the state over 181 elementary schools (cf. Ennis, 2007; McKinney, 2007; Savaria, 2008) and 83 secondary schools (Todd, 2010) for the period 2001-2005, more so even than the effect of free and reduced lunch. That same trend is apparent in statewide population means for the Comprehensive Test of Basic Skills third grade mathematics NCE achievement scores for the years 2001-2005, from \( \mu = 54 \) to \( \mu = 59.3 \). Collectively (see Analysis, Research Question 2), the America’s Choice schools did not share in this long term statewide progress.

Table 14 in Chapter IV provides further evidence on the effects of beginning year for the America’s Choice Schools, but in the opposite direction of that predicted by gradual increase over time in Kentucky schools. Consistent with the trend of continual improvement under KERA, steady statewide increases are shown for the state schools (as measured by percentage change compared to the year prior to implementation). That same pattern is shown for America’s Choice Schools C, G, H, and I. But for Schools D, E, and F (which began implementation in 2003 rather than 2002), the result was decline rather than progress. This outcome suggests that school to school differences, or perhaps
programmatic differences for the schools that began a year later (same intervention program but possibly differences or changes in consultants and staff along with other school initiatives) are a better explanation for these seven America's Choice Schools. Again, case studies and/or level of use data would be necessary to understand these implementation-outcome interactions.

Research Question 5

What is the effect of controlling for demographic factors for Research Questions 2-4?

The analysis for Research Question 5 utilized hierarchical multiple regression which incorporated the independent variables in sequential blocks to establish the prior effect of the demographic factors (size and percent free/reduced lunch) before introducing the implementation model blocks from Research Question 4. As in RQ4, the Comprehensive Test of Basic Skills (CTBS/5) Normal Curve Equivalent (NCE) third grade math scores for the seven schools served as the dependent variable. The steps for introducing variables were Step 1 (demographic variables--size, percent free/reduced lunch), Step 2 (years of implementation), Step 3 (differences between schools), and Step 4 (beginning year of implementation). As part of the analysis, the researcher noted that School I (2004 n = 12) did not report free and reduced lunch data for third grade for the year 2004. However, the small number of students involved had minimal impact on the overall results.

For Step 1, entering the demographic factors produced a non-significant model. In Step 2, year of mathematics implementation (Year 1, Year 2, and Continuation Year) was added to the equation while controlling for demographic factors. The model was significant, but the effect size was less than .01. Step 3 of the regression model added
differences between schools to the equation in addition to the demographic factors and year of implementation. For this step, the model was significant but the Adjusted $R^2$ was only .01, less than the minimal .02 change for large samples required for a substantive finding (Cohen et al., 2003). Although all of the schools participating in this study were low performing schools with a high percentage of students participating in the free/reduced lunch program, there was considerable variation among them. Schools F and I differed significantly from School C (the referent dummy coding category). Despite somewhat similar demographics, these schools did not all fare equally well during the America’s Choice project, but the differences had no practical effect on the outcomes.

Finally, Step 4 added beginning year of implementation to the equation. The year 2002 was the beginning year for schools C, G, H, and I while 2003 served as the beginning year for schools D, E, and F. The model was significant but the effect size remained unchanged at .01.

**Analysis**

For this analysis, discussion is limited to Step 1 of the 4 Step Hierarchical regression reported in Table 11. Step 1 addressed the effects of the demographic factors on third grade mathematics achievement. Steps 2-4 essentially mirror Table 10 (RQs 2-4) above and are not repeated here.

The seven Kentucky elementary schools that participated in the mathematics portion of the America’s Choice comprehensive reform model were similar demographically (see Table 2 in Chapter III). For this study, only free/reduced lunch and size were utilized as demographic controls.

Despite the fact that only four of the schools were Title I identified, all were generally low income and low achieving. Based on 2000 data, all schools except E (66%)
and F (38%) had free/reduced lunch participation varying from 93-98 percent (see Table 2 in Chapter III). However, the third grade only percentages for free/reduced lunch were considerably lower (see Table 6 in Chapter IV). It is not clear why the discrepancy in percent free/reduced lunch between Table 2 in Chapter III and Table 6 here. Cohort variation seems an unlikely explanation because the mean values in Table 6 represent four separate cohorts. One possibility is that the schools were becoming less impacted by poverty since 2000, but that seems inconsistent with the economic conditions in rural Kentucky.

Missing data is perhaps a better possibility. The analyses for Table 10 (no demographic variables) had $N = 1444$; Table 11 (with demographic factors added) had $N = 1032$. Given the well documented problems with measurement error for free/reduced lunch (see Harwell & LeBeau, 2010), the missing cases could well have been students who were eligible for but had no record of federal lunch participation. Restrictions of range and level of measurement are other possibilities. With generally low income schools, the range of variation across students’ lunch participation may not be sufficient to capture a relationship with math achievement. The actual coding of the free/reduced lunch variable (nominal) also leads to reduced power (in or out of program rather than a continuum of income levels).

Whatever the reason, the remaining cases demonstrated no significant relationship with third grade CTBS math scores, for either free/reduced lunch or size. Thus these results are inconsistent with the extensive base that indicates certain demographic factors play an important role in academic achievement. Percent free and reduced lunch or socioeconomic status is one of the most powerful (Coleman et al., 1966; S. K. Miller et al., 2006; Rothstein, 2004; Tate, 1997; Wilson, 1987). While the
achievement gap for race and income continues to be a concern for educators, these factors were not germane to this study. (For income, see discussion in preceding paragraphs; for race, the black-white achievement gap [cf. Jencks & Phillips, 1998] is not relevant to this study because these rural schools are essentially all white.) The two exceptions to this among the nine America’s Choice Schools were A and B, neither of which had full implementation of the math component, although both were included in the reading analysis (see VanMeter, 2005).

Kentucky provides additional support for schools with high populations of low socioeconomic students. Through Title I funding many schools adopt comprehensive reform models as a means of addressing achievement gap issues. The seven schools in this current study selected America’s Choice as a means of helping these students. Ross et al. (2004) found that when examining the effectiveness of implementing any comprehensive reform, demographic factors should be considered along with implementation factors, including degree and extent of fidelity to the practices that constitute a given model of reform.

It should be noted that three of the schools (D, E, and F) were not Title I. A policy related to the comprehensive reform models adopted by the Tier 3 Assistance schools in Kentucky is the funding source. The Kentucky Department of Education (KDE) identifies schools that are persistently low achieving (PLA) schools. These may or may not be schoolwide Title I schools that would have access to Title I funding. They may, however, be in need of assistance because of their Tier status. Therefore, these schools were able to use other funding sources, unknown to the researcher, to support the America’s Choice program as their method to address their low performance.

On the other hand, the demographic factors in the current study can be reviewed
as contextual background (generally low income, white, rural, low achieving schools) for which both educators and policy makers should be cognizant regarding any suggestions/strategies for improvement (cf. Carlson, 2003; Harwell & LeBeau, 2010; Lee, 2006). Yet, statistically, it is not surprising that the demographic factors in this investigation had little effect, due to restriction of range. Compared to the entire state, the range of these variables is both small and generally homogenous. This is true for both free/reduced lunch as a poverty indicator and size. This issue is even more problematic for ethnicity, Title I identification, rural-urban status, and English language learners--the four variables from Table 2 that were excluded from this study because of their restricted range. Accordingly, generalizability of these results is limited. The lack of statistical effect from such variables in this study should not suggest these variables do not still have important effects in more representative samples.

**Research Question 6**

For the three years of the America’s Choice mathematics program (Year 1, Year 2, Continuation 1), how does the progress of the America’s Choice schools compare to that statewide for the same period?

Two approaches were used for the analysis of Research Question 6. To look at the progress of the seven America’s Choice schools, the two cohorts based on starting dates were examined separately in comparison to the state for the three years of implementation (Y1, Y2, and C1). Schools that began the first year (C, G, H, and I) were compared to the state scores from 2002 through 2004, while those beginning in the second year (Schools D, E, and F) were compared to scores from 2003 through 2005.

**Approach One: One Sample t Tests**

In the first approach, one sample t tests were calculated (see Table 12 in Chapter IV). The results indicate that Schools C, G-I were significantly lower than the state for
Year 1 while Schools D-F were lower but not significantly so. Both cohorts were lower but the differences were not significant for Year 2; for the Continuation Year both groups were significantly lower than the state. Thus, for the overall set of one sample t tests, the America’s Choice schools were always below the state mean but for the two years of active implementation, the gap was only significant for Schools C, G, H, and I for Year 1. As the intervention progressed, however, the America’s Choice schools fell further behind the state so that by the end of the Continuation Year (with no active funding support), the gap had become significant and substantial.

Approach Two: %Change Scores

The second approach required multiple steps to calculate and then compare percentage change scores for the two cohorts of America’s Choice schools and the state. Step one included the calculation of mean math scores for the two cohorts of America’s Choice schools and for the state based on the staggered start dates, for each of the four years (see Table 13 in Chapter IV). Step two involved calculating unweighted percentage change scores (see Table 14 in Chapter IV) while step three included weighted percentage change scores. Both calculations were utilized because unweighted looked at the two cohorts of schools separately and weighted combined the cohorts into one group to examine America’s Choices schools as a whole. While the results from the weighted calculations responded to the research question directly, it masked the differences between the two cohorts. Both analyses were needed to provide a clear response to the question. Table 15 in Chapter IV provides the weighted percentage change scores for the seven America’s Choice schools combined and for the state combined for the implementation years (Year 1, Year 2, and Continuation Year).

In examining the findings from both approaches, the results for the two cohorts
differed slightly. When combining both cohorts, the overall results of the America’s Choice school reform project reflect steady progress during the two years of active implementation of the program (a bit greater than the state), followed by a precipitous decline during the continuation year when active support was removed, to a level below the baseline of scores for the year prior to implementation. (In contrast, the State continued its steady upward progress during that same year.)

*Analysis*

The Commonwealth of Kentucky is committed to the concept of excellence in education, embracing the idea that all children can learn, as evidenced through the enactment of KERA. Overall, Kentucky students have raised their achievement scores on the yearly Commonwealth Accountability Testing System (CATS) and the National Assessment of Educational Progress (NAEP) tests (Petrosko, 2000; Poggio, 2000).

The findings from this study provide additional evidence for statewide increase in achievement levels on accountability assessments. Weighted percentage change scores combining third grade CTBS mathematics NCE scores for the years 2001 to 2004 and for 2002 to 2005 demonstrate the steady progress. For Year 1, Year 2, and Continuation Year (the years of America’s Choice comprehensive school reform intervention—as compared to the year prior to the intervention), Kentucky posted percentage change gains of 2.73, 5.86, and 7.98. Clearly statewide mathematics scores on the nationally normed CTBS test were improving for the elementary school level.

In contrast, the current findings indicate that America’s Choice had varying degrees of impact on Kentucky schools implementing the mathematics component in third grade. Some schools (specifically C, G, H, and I) demonstrated steady improvement in achievement during the two years of supported implementation; these schools fell back
some during the Continuation year with no support, but still progressed compared to their beginning point in the prior year. In contrast, Schools D-F made slight progress for Year 1 and Year 2 but declined badly during the Continuation Year (almost 7% below their benchmark for the Prior Year.

During the implementation years (Y1 and Y2) of active support, schools had external support from the America’s Choice team in the form of school visits to provide support for the leadership team and training academies for the school level mathematics and literacy coaches. During this time, schools also developed internal levels of support for implementation of the various components of America’s Choice including team meetings, model classrooms, and training from their coaches. During the Continuation year (C1), the external level of support ended and schools were left to continue on their own with only their internal levels of support.

Noteworthy in this regard is the fact that Schools D-F did not “continue” with their implementation of the America’s Choice comprehensive reform model during the continuation year (see Table 1). These three schools did not begin the project until the second year and may have been less committed to seeing it through to the end. Certainly this result is consistent with that interpretation, although, again, level of use data or case studies would be needed to confirm this.

Although all of the America’s Choice schools continued to lag behind the state the outcomes differed precipitously for the two groups. The implication here is that what happens with comprehensive reform depends on what/how the faculty in a school implement that program. The same America’s Choice program adopted by schools with similar demographic backgrounds yielded noticeably different results. This suggests that schools must develop a strong understanding of the components of America’s Choice in
order to implement them fully and correctly. This must take place in order to see a positive impact on achievement scores.

In the literature, America's Choice has been shown previously to have a positive impact on achievement scores (May & Supovitz, 2006; Supovitz & May, 2004a, 2004b; Supovitz et al., 2002). Other research also indicated significant improvements on student scores on state level mathematics assessments during the implementation years of comprehensive reform models (Epstein, 2005; Sterbinsky et al., 2003). Thus despite the lack of progress generally in the current study, comprehensive school reform can and has produced positive results.

The current study did not look at level of use data; case studies to examine how precisely the America's Choice program was developed in each different school are not available. Such data would be necessary to understand why the two groups of schools had such different results. VanMeter (2005) studied reading achievement in the same schools and found differing results from school to school. Examining the level of program component implementation in each school was beyond the scope of work for each of these studies.

However, the literature on change (e.g., Ross & Gill, 2004; Sterbinsky et al., 2003; Supovitz & May, 2004b; Zhang et al., 2006) and on comprehensive school reform (see McCaslin et al., 2006; Supovitz & May, 2004b; VanMeter, 2005; Wetherill & Applefield, 2005) strongly supports the notion that what faculty do in response to comprehensive reform is the key factor in determining whether improvement occurs. Any comprehensive reform package provides a framework for school improvement. But schools are inhabited by real people. What and how they respond--changes (or not) in values, attitudes, instructional strategies, skill levels, etc.--are ultimately related to whether the school
actually improves.

One final note can be added here. Table 15 and Figure 4 present the combined weighted %Change scores for the seven America’s Choice schools compared to the State for the two years of active implementation and the continuation year without external support. The State made steady progress over the three years. America’s Choice improved during the two years of active programming (slightly ahead of the State’s pace), but then declined to a level below the baseline from the year prior to the beginning of the program. For all the literature on how difficult change is, the AC schools improved during their two years of mathematics implementation at a level equivalent to or greater than the state gains (change is hard but possible). The real challenge is institutionalization, i.e., sustaining improvements over time. Here these America’s Choice schools failed miserably (institutionalization is really hard, and we haven’t figured it out). The author is unaware of a similar decline of such precipitous proportions in the literature on comprehensive reform. It is almost as if the faculties of these schools let out a collective sigh of relief that the program was “finished” after the two years so they could rest, relax, and get back to business as usual (whatever they were doing prior to America’s Choice). The intervention would seem to have been viewed as a short-term interruption of norms that were focused on anything but continuous improvement. This contrasts with the statewide focus on continuous improvement to meet the goal of Proficient (a school score of 100 on the state’s accountability system). That long-term focus appears to have had a greater impetus toward improving achievement than a specific program with bounded intervention and no long term thrust.

The analysis in the previous paragraph needs to be put within the context of school to school variation. While the overall trends just discussed are clearly relevant, there were
differences from one school to the next. More importantly, contextual considerations in the overall intervention were also important. The three schools that began the America’s Choice comprehensive reform package in 2002 (a year after the initial group) had lower overall rates of improvement and had a much more precipitous drop during the sustainability year than the first group of four schools. Whether commitment was less or some other factor, clearly the tenor of expectations and norms of improvement would seem to have been less for the second group of schools. (Level of Use data were not analyzed to verify this.) Those who would consider change would do well to remember that intangibles such as effort, commitment, and long-term goals can and do make a difference in outcomes for schools and the students who reside within them.

Recommendations

Based upon the findings from this study, the researcher included recommendations for educators and others in the areas of policy, practice, and future research. Each is addressed in the sections that follow.

Policy

Each year Comprehensive Test of Basic Skills (CTBS/5) scores are disseminated in Kentucky elementary schools and decisions are made about success and instructional practices. Schools with low scores are required to undergo a Scholastic Audit using the Standards and Indicators for School Improvement (SISI) document (Kentucky Department of Education, 2004a). These standards and indicators represent a research-based means of evaluating the educational programs as they are being implemented in schools. These results constitute a road map to school improvement needs and goals.

The results of this study represent a partial program evaluation of one of
Kentucky’s efforts to help schools which were designated as Tier 3 Assistance on the Scholastic Audit. These schools were encouraged to adopt one of several comprehensive school reform packages to guide their improvement efforts. (The schools could choose between the following programs: Accelerated Schools Project, America's Choice, ATLAS Communities, Early Intervention in Reading, First Steps, Modern Red Schoolhouse, School Development Program, and Success for All.) This study examined the seven (of nine) schools that chose America’s Choice for which mathematics data were available.

First, the results of this study provide significant insight into decision making for school administrators considering adoption of America’s Choice comprehensive reform. The approximate cost for implementing America’s Choice is $72,000 per year for a minimum of three years. This cost would be one factor that, when added to the evaluation of the program and available research, would provide useful information to administrators involved in the decision making process. Time commitment is another factor. America’s Choice is an intense, involved comprehensive reform model that requires a large time commitment from all staff involved in order to achieve the desired outcomes. Level of commitment was beyond this study but may have provided insight into the individual school variations in the results.

Second, it is imperative that teachers and administrators make the most of any comprehensive reform model they elect to adopt by understanding and implementing all of its essential components. The program selected may not matter as much as the way in which it is implemented. Too often schools are looking for the golden bullet and this study indicates that programs vary in their effectiveness. The school to school differences in progress suggest that quality and consistency of implementation, not the specific program, is what matters most.
Third, the Commonwealth of Kentucky offers a school-wide improvement model consisting of the *Standards and Indicators for School Improvement* (SISI) and the Scholastic Audit to assist struggling schools with their academic achievement efforts. Schools should take full advantage of the specific data on their strengths and weaknesses provided through these resources. This can help guide schools in the process of selecting an appropriate comprehensive reform model to fit their educational needs most effectively. School leaders must look not only at the data itself which tells what is happening but also include qualitative monitoring to provide further explanation in regards to the why this is happening.

**Practice**

The National Council of Teachers of Mathematics (NCTM) has been a leader in providing guidance to shape the mathematics reform movement. The *Principles and Standards for School Mathematics* (NCTM, 2000) provides the framework for the development of strong math programs that move away from the traditional approach and focus on a more constructivist teaching model that emphasizes development of critical thinking skills, problem solving, and making real life connections. The math component of America’s Choice focuses on developing basic skills through problem solving, critical thinking, and conceptual understanding. This overlap between the National Council of Teachers of Mathematics (NCTM) *Principles and Standards* and America’s Choice suggests that implementation of this comprehensive reform package should result in improved mathematical instruction and student outcomes consistent with the literature on the effects of implementing the NCTM approach.

Yet, overall, the results for the seven America’s Choice schools that implemented math in Kentucky were disappointing, essentially no change from where they started. The
two cohorts of schools (successive years for stating America’s Choice) had different results even though all schools were similar demographically. These differences across schools suggest that the math component of America’s Choice is differentially effective, depending on what happens in each school.

The current study did not analyze data on how the instruction was conducted (level of use of the intended intervention). That suggests the need for more research on actual implementation practices (see Future Research below).

But clearly, based on the results of this study, the schools achieved differentially well. This implies that the focus for practice should be not on the program being implemented, but rather, what teachers do during instruction. If outcomes for math are to change, math instruction must change (Davis et al., 1990; Elliott, 1996; Gregg, 1995).

Given the literature on resistance to changing instructional practices (cf. Cuban, 1990; Henry & Clements, 1999) perhaps the focus for schools should be less on an overall comprehensive model and more on instructional practices. It is certainly possible that the emphasis on America’s Choice (or other models) displaces teacher efforts to change their instruction. It is much more likely that teachers in the seven schools thought of themselves as “doing America’s Choice” than they did as “changing my instruction.” The former is a relatively safe school program. The latter is personally threatening. Yet, ultimately, schools do not change unless individuals do. There needs to be a group dynamic (cf. Brookover et al., 1982) but there must also be personal responsibility for effective change.

Perhaps even more relevant is the fact that these schools did make progress during the two years of active implementation. It was during the continuation years (trying to sustain what they had achieved) that these America’s Choice schools faltered. Once again, this suggests that institutionalization is even more difficult than initial change and that
future policy and practice need to emphasize this reality for the teachers and principals who populate Kentucky’s (and America’s) schools

*Future Research*

This study focused on examining the impact of implementing the mathematics component of America’s Choice comprehensive reform in seven Kentucky elementary schools. The researcher compared progress of the America’s Choice Schools to the state and then used multiple regression to examine the relationships between implementing the mathematics component of America’s Choice and student achievement scores in math. While the results of this study provide insight to assist administrators in consideration of implementing America’s Choice, further research is needed to examine other areas of this comprehensive reform model. Suggestions for future research should focus on looking at America’s Choice comprehensive reform through different perspectives.

First, future research could compare a cohort of America’s Choice schools with demographically similar schools to determine the impact of implementation. The researcher could compare any one of the components (reading, writing, mathematics, or science) on the impact of school wide implementation.

Second, a study could follow the *same* students within a school implementing America’s Choice reform over a period of time, looking at yearly progress in one or more areas of America’s Choice component implementation. This type of study would provide information about the impact on student achievement over time, as opposed to the current study which examined successive cohorts.

Third, since research has been conducted in Kentucky schools with a focus on the reading and mathematics components of America’s Choice, future studies could be conducted examining the writing component. The focus could be on one grade level
(a cohort model) or follow students over a period of years (longitudinal) to track their
achievement in writing. The content area of science could also be examined in the same
way.

Fourth, this study was quantitative in nature and did not include any measures of
levels of implementation of the program in any schools. This information was available
but was beyond the scope of this research. Future qualitative research could be conducted
with a focus on the involvement of the leadership team measuring the levels of
implementation, including teacher and student survey data, to determine the impact of
implementation on student achievement scores. Including this information along with
analyzing student achievement data would provide more information about
implementation, including why the considerable variation from school to school.

Fifth, future research could include a qualitative case study that looks at one
particular school and documents the change process over time from when they begin the
program and continue through the final year when the school receives no direct support
for implementation of the program. This could include teacher interviews and
observations to capture detailed information regarding the degree of implementation of
the program components in all areas.

Sixth, future research could include conducting an analysis of teachers and
their instructional practices to attempt to establish a relationship between teacher
effectiveness and student achievement based on implementation of the America’s Choice
comprehensive reform components. This study could focus on any or all of the content
areas addressed in America’s Choice and would provide data regarding teacher differences
in instruction and the impact on outcomes. As part of this analysis, the instruction could be
evaluated on a framework of best practices for the particular content area. For the current
study, this would be based on the *Professional Standards for Teaching Mathematics* (NCTM, 1995) and *Standards for School Mathematics* (NCTM, 2000b).

Seventh, future research could be conducted following the same group of schools excluding School F since it was an outlier. Of the seven schools in this study, School F was the only school not in Tier 3 Assistance implementing America’s Choice. The results were unsteady and it would be interesting to view the analysis with the inclusion of School F.

Eighth, research could be conducted as an extension to this study looking at the impact of absenteeism on mathematics achievement. The schools in the current study were all low socio-economic schools with varying rates of absenteeism and teacher turnover. This demographic variable could be included to determine its impact, if any, on mathematics achievement.

Finally, the trend for both cohorts of schools in this study (Schools C, G-I beginning 2001; Schools D-F beginning 2002) was improvement followed by a decline during the Continuation Year when active support was no longer available. (The amount of progress and then decline was very different for the two groups, but both demonstrated the same phenomenon.) In contrast, the progress by the state overall was steady, each year reflecting higher achievement than the year before. One interpretation of this is that the America’s Choice schools (Tier 3 Assistance in Kentucky’s system of evaluation--the lowest level mandated to do “something”) did their two years of “penance” and then relaxed because they had “finished” the intervention. In contrast, statewide, the schools did not relax after two years because they were not in a “program”; rather, they were thinking in terms of the long term goals for continual improvement (value-added progress--cf. S. K. Miller, 1992) under the KERA reforms. Each year of effort and improvement is simply one more step towards a demanding goal: proficient for the school by 2014. (Kentucky’s
criterion referenced accountability standards are pegged to the NAEP, considered to be “world class”--cf. Rothstein, 2004.)

This “we’re done with this program! Phew! Time to kick back!” phenomenon could be related to the amount of additional time and effort that is required as part of America’s Choice. Since teachers knew from the beginning of implementation that this was a short term commitment, they may have viewed the program differently. Further investigation of this possibility (in both existing studies as well as new databases) could perhaps produce greater insight about why intervention programs are so often successful in the short term but fade in the long run, one more instance of “this too did pass.”

Conclusions

The Kentucky educational system has high standards and expectations for all students in the schools across the Commonwealth. The goal is for all students to reach Proficiency (a composite score of 100 out of 140 calculated from CATS scores, the Academic Index, and non-academic factors) by 2014. Accountability and instruction guide teachers in reaching this high goal. As schools move along their individual path, it is apparent that some schools are falling behind and struggling.

Kentucky has developed a support system to assist these schools consisting of the Standards and Indicators for School Improvement: Kentucky’s Model for Whole School Improvement (SISI) and the Scholastic Audit process (Kentucky Department of Education, 2004a). As a result of using these tools, schools that are in Tier 3 Assistance can utilize the information to guide them in the selection of an appropriate comprehensive reform design model to address their needs.

Comprehensive School Reform (CSR) designs provide a model for schools that need to improve their educational programs in an effort to increase student achievement.
Kentucky schools that were among those making the least progress toward the state goal of Proficient by 2014 had the option and funding support to select and implement a Comprehensive Reform Model. Nine Kentucky elementary schools chose to implement America’s Choice comprehensive reform to help improve their academic achievement.

VanMeter (2005) studied the impact of America’s Choice reform on third grade reading achievement in these nine Kentucky elementary schools. This study builds upon the work of VanMeter by examining the math component of America’s Choice in seven of the nine schools implementing the math component during all four years of the study. The guiding question was: What are the effects of implementing America’s Choice comprehensive school reform design in mathematics in seven elementary schools in Kentucky?

In order to understand the effects of implementing the math component of America’s Choice the Comprehensive Test of Basic Skills (CTBS/5) Normal Curve Equivalent (NCE) mathematics scores for third grade students were examined for year prior to implementation (P1), two years of supported implementation (Y1 and Y2), and the continuation year (C1) with no direct support. Due to the small number of schools selecting America’s Choice and the “forced to volunteer” status of being designated as a Tier 3 Assistance school, the results of this study must be generalized with caution.

Six research questions guided the analysis by examining the central research question from several perspectives, all of which led to and supported a similar conclusion. Although the seven Kentucky elementary schools were able to demonstrate some impact during the active implementation of this model, as a whole, they were unable to sustain its impact in the year following when supportive resources were no longer available. Not only did they lose ground in the last year, three of the schools actually slid below (and
considerably so) their starting point for the Prior Year.

There were two managerial implications for this study. First, the sample included predominately low socio-economic schools wherein research shows a generally high degree of student mobility (Guskey, 1997; Harwell & LeBeau, 2010). Not all of the students who took the Comprehensive Test of Basic Skills (CTBS/5) mathematics assessment at the end of each year necessarily attended the school for the entire year to receive the effects of the implementation of the math component. (This could be related to the very high number of students with missing data as discussed above.) Secondly, low socio-economic schools tend to have high teacher turn-over rate. This impacts the quality and consistency of the math instruction throughout the year. These factors could have impacted the results of this study. Unfortunately, neither of these possibilities was addressed directly in the current study. It would require further research to examine these eventualities. (Teacher turnover could be examined over the course of the intervention in case studies. Student mobility data could be obtained from the Kentucky Performance Report, an archive of demographic and achievement data from Kentucky schools. It is plausible, at least, that student mobility would be a factor in the missing data noted for the much smaller N in Table 11 (with demographic variables included) versus Table 10 (achievement data only).

The evidence from this study added to the broader research literature on the impact of implementing America’s Choice varying from school to school. This could prove useful for school leadership teams in the role of analyzing and selecting an appropriate comprehensive reform model that best meets their needs. Previous investigations of America’s Choice have examined teacher implementation (Supovitz & May, 2004a), and reading achievement (Supovitz & May, 2004b; VanMeter, 2005).
Because the current study constitutes a conceptual replication of the VanMeter (2005) investigation, it is important to be explicit about how it differs. Two major distinctions are obvious: (a) VanMeter examined Reading whereas this study looked at mathematics; (b) the current study was limited to the seven schools (of the nine in the VanMeter study) that had complete data for mathematics. This result is related to the planned sequence of change for America’s Choice: reading in the first year followed by math in the second year. The other subjects (writing and science) were begun in year one and three, respectively, but were not addressed by either VanMeter or the current work. Apart from this overall distinction, the current study represents an advance over the VanMeter (2005) investigation with respect to the actual research questions and analyses that were conducted. VanMeter conducted an exploratory study utilizing a series of t tests to analyze the data. The current study went further to address the six research questions: utilizing a cohort model with hierarchical regression to examine the separate effects of year of implementation, differences across schools, and beginning year of the program (because the state as a whole was progressing), plus t tests and graphic presentation to compare the America’s Choice Schools to the state regarding rate of overall progress on school improvement.

Mathematics reform initiatives focus on shifting instructional emphasis towards problem solving, reasoning, and making connections between topics and concepts by providing opportunities for students to develop communication skills (see Carpenter & Moser, 1984; Riordan & Noyce, 2001). The mathematics workshop format of America’s Choice is structured in such a way that students have the opportunity to develop these skills. The current study focused on America’s Choice, specifically the mathematics component and its impact on student achievement. The findings add to the body of
knowledge for math instruction and comprehensive reform models by providing insight into the impact on low socio-economic elementary schools seeking to improve mathematics achievement.

The results here support the use of best-practices in math as an approach to school improvement (the math scores improved for all seven of these schools during the two years of active support). However, the study also provides a cautionary tale of how difficult positive change is to sustain. All seven schools regressed in the year following active support/implementation of the America’s Choice comprehensive intervention and one group fell to a level below their baseline from the year prior to the project.

Title I funds continue to be spent on comprehensive reform models, such as America’s Choice, to support schools in their quest to increase student achievement scores in reading and mathematics as they move towards reaching their goals set for the year 2014. It is the responsibility of administrators and educators to select and fully implement all elements and components of the selected comprehensive reform model if schools are to achieve the expected outcome of increased student achievement to prepare students to become successful, effective members of a global society. This study, and the literature on comprehensive school reform generally, provide evidence on how difficult that is. It requires changing the culture of the school as well as the instructional practices. Simply adopting a reform model is not enough. The real work is helping teachers and administrators change their practices, and then supporting them to sustain those changes over time. (This study suggests that institutionalizing change is considerably more difficult than short term modification during active program implementation.) The students in Kentucky’s schools are waiting for this transformation.

176
REFERENCES


Beam, M., & Faddis, B. (2001). The implementation of the comprehensive school reform demonstration program: The focus and progress of schools' work in seven Midwest states. Naperville, IL: North Central Regional Educational Laboratory.


social systems and student achievement: Schools can make a difference. New York: Praeger Publishers.


Mifflin Co.


APPENDIX A

AFFIDAVIT OF NONDISCLOSURE REQUESTING DATA
AFFIDAVIT OF NONDISCLOSURE

Senior Outreach Specialist
WIDA Consortium
Wisconsin Center Educ. Research

(Date of Assignment of KDE Project) dissertation
PhD research for
3rd grade mathematics scores for Ky schools
Requested Kentucky Department of Education Date Base or File Containing Identifiable Information

201 S. High Point Rd. Apt: 308F Madison, WI 53717

(Address)

1, Sonia James Upton, do solemnly swear (or affirm) that when given access to the subject Kentucky Department of Education data base or file, I will not

(i) use or reveal any individually identifiable information furnished, acquired, retrieved or assembled by me or others for any purpose other than statistical purposes specified in the KDE survey, project, contract, or proposed research;

(ii) make any disclosure or publication whereby a sample unit or survey respondent could be identified or the data furnished by or related to any particular person can be identified; or

(iii) permit anyone other than the individuals authorized by the Commissioner of the Kentucky Department of Education to examine the individual reports.

Sonia James Upton
(Signature)

[The penalty for unlawful disclosure is a fine of not more than $250,000 (under 18 U.S.C. 3571) or imprisonment for not more than five years (under 18 U.S.C. 3559), or both. The word "swear" should be stricken out when a person elects to affirm the affidavit rather than to swear to it.]

City/County of Dane Commonwealth/State of Wisconsin
Sworn to and subscribed before me this 23rd day of March 2010. Witness my hand and official Seal

Teresa L. Mason
Notary Public/Seal

My Commission expires September 2, 2012

191
APPENDIX B

PERMISSION LETTER FROM IRB UNIVERSITY OF LOUISVILLE
This study was reviewed on 10/19/2010 and determined by the chair of the Institutional Review Board that the study is exempt according to 45 CFR 46.101(b) under category (4) Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects. The study is exempt only if information that could identify subjects is not recorded.

Since this study has been found to be exempt, no additional reporting, such as submission of Progress Reports for continuation reviews, is needed. If your research focus or activities change, please submit a Study Amendment Request Form to the IRB for review to ensure that the study still meets exempt status. Best wishes for a successful study. Please send all inquires and electronic revised/requested items to our office email address at hsppofc@gwise.louisville.edu.
APPENDIX C

PERMISSION LETTER FROM IRB WESTERN KENTUCKY UNIVERSITY
In future correspondence, please refer to HS11-078, October 28, 2010

Sonia James Upton  
c/o Dr. Stephen Miller  
EALR  
WKU

Sonia James Upton:

Your research project, Mathematics Achievement: The Impact of America's Choice in Kentucky Schools, was reviewed by the HSRB and it has been determined that risks to subjects are: (1) minimized and reasonable; and that (2) research procedures are consistent with a sound research design and do not expose the subjects to unnecessary risk. Reviewers determined that: (1) benefits to subjects are considered along with the importance of the topic and that outcomes are reasonable; (2) selection of subjects is equitable; and (3) the purposes of the research and the research setting is amenable to subjects' welfare and producing desired outcomes; that indications of coercion or prejudice are absent, and that participation is clearly voluntary.

1. In addition, the IRB found that you need to orient participants as follows: (1) signed informed consent is not required; (2) Provision is made for collecting, using and storing data in a manner that protects the safety and privacy of the subjects and the confidentiality of the data. (3) Appropriate safeguards are included to protect the rights and welfare of the subjects.

This project is therefore approved at the Exempt from Full Board Review Level.

2. Please note that the institution is not responsible for any actions regarding this protocol before approval. If you expand the project at a later date to use other instruments please re-apply. Copies of your request for human subjects review, your application, and this approval, are maintained in the Office of Sponsored Programs at the above address. Please report any changes to this approved protocol to this office. A Continuing Review protocol will be sent to you in the future to determine the status of the project. Also, please use the stamped approval forms to assure participants of compliance with The Office of Human Research Protections regulations.

Sincerely,

Paul J. Mooney, M.S.T.M.  
Compliance Coordinator  
Office of Research  
Western Kentucky University

cc: HS file number Upton HS11-078
Appendix D
Standards and Indicators for School Improvement

The Standards and Indicators for School Improvement represents the framework for schools to maximize potential for all students by delineating professional standards in these areas, organized into three broad categories--Academic Performance, Learning Environment, and Efficiency (KDE - Division of School Improvement, 2003). Each of the nine standards include a number of indicators, ranging from 5 to 16. The indicators are listed in KDE (2004), Standards and Indicators for School Improvement: A Kentucky Model for Student-Centered Accountability.

Academic Performance

Standard 1 (Curriculum): The school develops and implements a curriculum that is rigorous, intentional, and aligned to state and local standards.

Standard 2 (Classroom Evaluation/Assessment): The school uses multiple evaluations and assessment strategies to continuously monitor and modify instruction to meet student needs and support proficient student work.

Standard 3 (Instruction): The school’s instructional program actively engages all students by using effective, varied and research-based practices to improve student academic performance standards.

Learning Environment

Standard 4 (School Culture): The school/district functions as an effective learning community and supports a climate conducive to performance excellence.

Standard 5 (Student, Family and Community Support): The school/district works with families and community groups to remove barriers to learning in an effort to meet the
intellectual, social, career, and developmental needs of students.

Standard 6 (Professional Growth, Development and Evaluation): The school/district provides research-based, results driven professional development opportunities for staff and implements performance evaluation procedures in order to improve teaching and learning.

Efficiency

Standard 7 (Leadership): School/district instructional decisions focus on support for teaching and learning, organizational direction, high performance expectations, creating a learning culture, and developing leadership capacity.

Standard 8 (Organizational Structure and Resources): The organization of the school/district maximizes use of time, all available space and other resources to maximize teaching and learning and support high student and staff performances.

Standard 9 (Comprehensive and Effective Planning): The school/district develops, implements and evaluates a comprehensive school improvement plan that communicates a clear purpose, direction and action plan focused on teaching and learning.
APPENDIX E

PEARSON R CORRELATION MATRIX
Table E-1

Pearson r Correlation Matrix for NCE Mathematics Score and Demographic Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>SIZE</th>
<th>%F/RL</th>
<th>NCE Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIZE</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>%F/RL</td>
<td>-.029</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>NCE Math</td>
<td>.140</td>
<td>-.006</td>
<td>--</td>
</tr>
</tbody>
</table>

Note. SIZE = number of third grade students, %F/RL = Percent free and reduced lunch, NCE Math = Normal Curve Equivalent Math.

*p < .01.
CURRICULUM VITAE
Sonia James Upton
201 S. High Point Road Apt. 308F
Madison, Wisconsin 53717
supton@wisc.edu

Academic Degrees

PhD, Organizational Leadership and Administration
University of Louisville & Western Kentucky University
Louisville and Bowling Green, Kentucky
Completed May 2012
Dissertation: Mathematics Achievement: The Impact of America’s Choice in Kentucky’s Schools

English as a Second Language Endorsement K - 12
Western Kentucky University, Bowling Green, Kentucky

Master of Arts in Education, May 1987
Western Kentucky University, Bowling Green, Kentucky
Major: Early Childhood & Reading

Bachelor of Science in Education, May 1982
Western Kentucky University, Bowling Green, Kentucky
Major: Elementary Education K – 8
Minor: Dance and Theater

Certifications

Elementary Certificate for Instruction, Grades K – 8
English as a Second Language Endorsement PreK – 12
Certificate of Instruction Early Childhood

Employment

WIDA Consortium Policy Administrator
Wisconsin Center of Education Research,
University of Wisconsin – Madison, May 2011 to present
Prepare and disseminate policy briefs and reports for WIDA Board. Manage the development and disseminate Consortium-wide communication. Liaison between WIDA departments (professional development, operations, research, educator resources and technology) and Consortium member states. Facilitate focus groups, surveys, and technical assistance.
Senior Outreach Specialist – WIDA Consortium  
Wisconsin Center of Education Research,  
University of Wisconsin – Madison, January 2010 to May 2011  
Develop and design new trainings. Provide technical assistance to the 27 Consortium member states. Create web-based trainings. Develop Resource Guides for state use relating to various topics for English learners. Create and provide Data Analysis Trainings.

ELL Program Consultant – Title III/LEP & Immigrant,  
Kentucky Department of Education  
Frankfort, Kentucky June 2008 – January 2010  
Provide technical assistance to administrators and educators in grades PreK – 12 who provide instructional services for English language learners in Kentucky schools. Create, plan, organize and implement statewide professional development trainings. Design, organize and implement a monitoring system of English language learners and immigrant students’ district programs. Editor and publisher for quarterly ESL newsletter for the state. Monitor and update English language learner webpage at KDE.

English Language Learner (ELL) District Coach  
Bowling Green Independent Schools  
Bowling Green, Kentucky, 1997-2008  
Plan, organize and implement monthly professional development trainings related to best practice pedagogy, instructional strategies and ESL standards implementation for the district English as a Second Language teachers in elementary, middle school and high school. Plan and implement ELP ACCESS assessment trainings. Train and coach ESL teachers in successful use of the state ELP data recording system.

English Language Learner (ELL) Specialist/Coach  
Dishman-McGinnis Elementary,  
Bowling Green, Kentucky, 1997 – 2008  
Effectively served Dishman-McGinnis English language learners in Preschool through 5th grade. Implemented developmentally appropriate, research based best practice strategies in sheltered instruction classes. Effectively established collaborative instructional content lesson with classroom teachers. Coordinated all facets of the Extended School Services program. Additionally served as the school Technology Coordinator, DIBELS Assessment Coordinator, GRADE/GMADE Assessment Coordinator, mathematics coach.

Effectively designed and delivered professional development
trainings to school administrators, content area teachers and ESL instructors in school districts all across Kentucky based on individual district needs.

**Kindergarten Assistant**
**Dishman-McGinnis Elementary**
**Bowling Green, Kentucky, 1991-1997**
Successfully planned and implemented daily reading instruction to small groups. Managed the student data and record keeping system for the kindergarten class. Assisted the teacher with instruction and support services. Organized, planned and implemented family reading night programs quarterly. Taught reading and writing classes to the ESL students’ schoolwide.

**Professional Trainings and Affiliations**
- ACCESS for ELLs® Test Administrator
- CESA 7 Data Analysis Trainer
- DIBELS Assessment Trainer
- NCEE Math Institute Coach
- CHAMPS Training
- Consolidated Planning Committee Training
- Curriculum Mapping
- Steven Covey’s 4 Roles of Leadership
- Kentucky Reading Project-Model Classroom & Instructor Transformation Plan Training
- KDE - ELL Assessment evaluation/selection advisory team
- WIDA cut score review/setting for Kentucky
- KDE - WIDA ELP Standards alignment project team leader
- SIOP Trainer for Kentucky
- Critical Friends Group Certified Coach

**Publications**
- Co-author - “*Implementing the ESL Standards into Classroom Practice: K-2 Volume*, TESOL and CAL
- Author - *Impacting Diverse Learners through Literacy Instruction*, Kentucky Reading Journal, winter 2001
- Author - *Comparing Three Computer Based Programs for English Language Learners*, TESOL Matters, 2000
<table>
<thead>
<tr>
<th>Professional Organizations</th>
<th>TESOL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>KYTESOL-past president</td>
</tr>
<tr>
<td></td>
<td>IRA</td>
</tr>
<tr>
<td></td>
<td>Phi Delta Kappa</td>
</tr>
<tr>
<td></td>
<td>ASCD</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>National Conference Presentations</th>
<th>NCEE National Conference, Orlando, Florida</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TESOL National Conference, New York</td>
</tr>
<tr>
<td></td>
<td>TESOL National Conference, St. Louis, Missouri</td>
</tr>
<tr>
<td></td>
<td>International Reading Association National Conference</td>
</tr>
<tr>
<td></td>
<td>Southeast TESOL Regional Conference</td>
</tr>
<tr>
<td></td>
<td>NAYEC National Conference</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>State Conference Presentations</th>
<th>KYTESOL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Kentucky Reading Association</td>
</tr>
<tr>
<td></td>
<td>Kentucky Teaching and Learning Conference</td>
</tr>
<tr>
<td></td>
<td>Early Childhood Education Conference</td>
</tr>
<tr>
<td></td>
<td>Headstart/Early Childhood Conference</td>
</tr>
<tr>
<td></td>
<td>Migrant Education Annual Conference</td>
</tr>
<tr>
<td></td>
<td>Kentucky Reading Project</td>
</tr>
</tbody>
</table>

| Research Focus                  | English language learners, language acquisition, academic language, student achievement, Response to Intervention |