6-20-2011

Comparison of Secondary Principals' Use of Data Systems to IncreaseStudent Achievement in Mathematics as Measured by Standardized Assessments

Joshua Williams
Indiana University of Pennsylvania

A COMPARISON OF SECONDARY PRINCIPALS' USE OF DATA SYSTEMS TO INCREASE STUDENT ACHIEVEMENT IN MATHEMATICS AS MEASURED BY STANDARDIZED ASSESSMENTS

A Dissertation

Submitted to the School of Graduate Studies and Research

in Partial Fulfillment of the

Requirements for the Degree

Doctor of Education

Joshua Williams

Indiana University of Pennsylvania

May 2011

Indiana University of Pennsylvania
The School of Graduate Studies and Research
Department of Professional Studies in Education

We hereby approve the dissertation of

Joshua D. Williams

Candidate for the degree of Doctor of Education

Sue A. Rieg, Ed.D.
Professor of Education, Advisor

Robert E. Millward, Ph.D.
Professor of Education

Cathy C. Kaufman, Ph.D.
Professor of Education

ACCEPTED

Timothy P. Mack, Ph.D.
Dean
School of Graduate Studies and Research

Title: A Comparison of Secondary Principals' Use of Data Systems to Increase Student Achievement in Mathematics as Measured by Standardized Assessments

Author: Joshua D. Williams

Dissertation Chair: Dr. Sue Rieg

Dissertation Committee Members: Dr. Robert Millward
Dr. Cathy Kaufman

The pressure to meet the demands of *No Child Left Behind* (NCLB) *Act* coupled with poor results by secondary students on national assessments in mathematics have forced school principals to develop skill sets in the use of data in efforts to increase student performance on standardized assessments. The effective use of data by school principals has been shown to improve student performance. This study compared the use of data systems by secondary principals, and their perceptions of data type and data tool effectiveness in enhancing student achievement in mathematics in high and low performing schools.

One-hundred forty-seven secondary school principals in high and low performing schools, as measured by the percentage of grade 11 students rated proficient on the 2008-2009 PSSA standardized mathematics assessments, were surveyed to examine what principals in high performing schools do differently in their use of data to increase student performance in mathematics as measured by standardized assessments. Analysis of the survey focused on 11 sub-categories that included: use of input data; use of process data; use of outcomes data; use of satisfaction data; perceived effectiveness of input data; perceived effectiveness of process data; perceived effectiveness of outcomes data; perceived effectiveness of satisfaction data; use of data systems; use of data tools; and, perceived effectiveness of data tools used.

Significant differences were found between principals in high and low performing schools in their use of data systems and overall use of data tools. Principals in lower performing schools were found to use data systems to a greater extent and reported a significantly higher use of data tools in their efforts to improve mathematics scores. Chi-square analysis showed some significant differences between principal groups in the use of benchmark assessments and related data tools. The researcher concluded that principals in low performing schools see the use of data as the cure for poor student achievement in mathematics as they feel the pressure of high stakes accountability under NCLB. Pressure to increase student achievement may have resulted in greater use of data tools and systems as principals in low performing schools search for solutions to poor student achievement.

ACKNOWLEDGEMENTS

I am grateful for the encouragement and insights of my dissertation chairperson, Dr. Sue Rieg. Her feedback was invaluable, constructive, and timely throughout this process. My thanks to Dr. Cathy Kaufman who, as one of my dissertation committee members, provided insights that were helpful in my work toward reaching this goal. Thanks to Dr. Bob Millward for his feedback as one of my dissertation committee members, and also for his leadership in running a program that has served me well in this process and in my professional career.

Special thanks to my wife, Renee. During this process you encouraged me when I needed it, you never complained, and were always there to listen. I would not have finished this without you. Thank you to my girls, Cassie, Kaley, and Lacey. Each of you has helped me to reach this goal in ways you do not realize. I am blessed to have such wonderful daughters.

Though he would not take credit for it, my father is an exceptional leader and an outstanding educator. The advice, counsel, and support he has provided me through this process and over my career is irreplaceable. Though I am not there yet, he has set the professional bar that I am working toward. He is and will always be the original Dr. Williams. Thanks to my mom for always believing in me and supporting me. Thank you to my brother who taught me early in my life that perseverance and toughness are important if one is to accomplish difficult things. This is a lesson that has helped me in this process and in my life more than he knows.

My thanks to Henry Shepard who built on what I learned from my brother. In every conversation, he told me he was proud and asked how close I was to finishing. It motivated me. Lastly, I want to thank my colleagues from Indiana University of Pennsylvania's administrative and leadership studies program (Cohort 6) for their support, especially Dr. Luke Lansberry who "turned the tide" for me in one conversation.

Table of Contents

List of Tables

Table | Page

Table Page

List of Figures

CHAPTER 1

STANDARDIZED ASSESSMENTS AND ACCOUNTABILITY

The *No Child Left Behind Act of 2001* (NCLB) was signed into law by President George W. Bush on January 8, 2002. The law dictated that every state receiving federal monies must demonstrate that it has a program of high quality annual student assessments, which can be used to determine the annual performance of every school and local education agency (No Child Left Behind Act, 2001). In fulfillment of the this legislation, the state of Pennsylvania conducts annual, standardized student assessments using a mechanism called the Pennsylvania System of School Assessment (PSSA). The PSSA is a standardized exam administered to all students in grades 3 through 8 and grade 11, and is designed to assess student progress toward proficiency on state standards for mathematics, reading, writing, and science (Pennsylvania Department of Education, 2003).

NCLB and assessments like the PSSA have spawned a thrust for the collection and analysis of data to support decision making in schools (Education Commission of the States, 2002). Assessment data in mathematics and reading is reviewed annually to assess individual school and district progress toward the 100% proficiency by 2014 requirement set by the NCLB (2001) for all students. These data for mathematics and reading assessments are reported as adequate yearly progress by the Pennsylvania Department of Education for districts, schools, and student sub-groups as the percentage of students that scored in the proficient range by subject. The average percentage of students in all grades assessed scoring proficient on PSSA exams has steadily increased over the past five years showing an increase in proficiency of 7.3% in mathematics and 6% in reading. Though assessment data in mathematics and reading indicates an increase in the percentage of students reaching proficiency, gaps exist between the average percentage of secondary students and students in grades third through eighth that score proficient on the PSSA reading and mathematics exam. While this gap has been significant for reading assessments (ballooning to 6% during the 2008-2009 school year), it has been far greater for mathematics ranging between 13% and 15.7% over the past five years including the largest gap in the 2008-2009 school year. Pennsylvania students' mathematics results support Schmidt's (1998) contention based on the Third International Mathematics and Science Study (TIMMS) that United States students' proficiency in mathematics declines from grade 4 to 12.

Background of the Study

Historical Perspectives

Standardized testing is not a novel idea for schools. This type of assessment has been used in schools for almost a century. In the mid to late 19th century, the influx of immigrants speaking many languages, the explosion of school enrollment, and an interest in applying industry efficiency techniques to education inspired the development of standardized testing (Pulliam & Van Patten, 1999). The common thinking of the time was that standardized testing would provide a means to classify and sort students to effectively educate them. Proponents of standardized testing at the time were critics of the American system of education including many prominent college and university administrators who felt schools were not preparing students well enough. Standardized testing progressed through the 20th century with the development of IQ testing and achievement testing, but the notion of school accountability based on standardized test data was formed from the outset (Pulliam & Van Patten, 1999).

Klein (2003) suggested that the history of mathematics education in the United States reached its most contentious point during the 1990s culminating in what was termed the "mathematics wars" by then U.S. Secretary of Education, Richard Riley. This conflict was not dissimilar to previous debates over education reform that occurred periodically throughout the 20th century. The root of the conflict over mathematics education stems from the influence of a progressivist approach to American education for most of the 20th century that was fostered by educational icons like William Heard Kilpatrick and John Dewey. Emphasis in American mathematics curricula has fluctuated between the progressivist leaning of a learner centered, discovery approach to mathematics that seeks to incorporate "real world" experiences, and more abstract mathematics curricula that emphasizes basic skills and facts often taught utilizing drill and instruction.

Though the progressivist approach to mathematics education was influential for the bulk of the 20th century, its support weakened in the face of educational "wake up calls" like the launching of the Soviet satellite Sputnik in 1957 and the compelling 1983 report, *A Nation at Risk*. The perceived loss in a race for space program dominance with the Soviets shattered the belief that the American education system was producing the best students in mathematics and science. In response to calls for reform, the *National Defense Education Act* was passed in 1958 aimed at increasing the number of mathematics, science, and foreign language majors while also contributing to school instruction. Mathematics education in the United States shifted to what was termed the "New Mathematics" program that for the first time included significant input from mathematicians and engineers in the planning of K-12 mathematics curricula while placing an emphasis on coherent logical explanations for the mathematical procedures taught in schools (Klein, 2003). By the early 1970s, the "New Mathematics" era was gone and the progressive approach to education returned with the Open Education Movement, which was met with a wave of opposition calling for a return to basic mathematics skills instruction in the mid-1970s.

American student mathematics scores evaluated during the 1970s with standardized basic skills assessments declined over the course of the decade bottoming out in the early 1980s.

The landmark report, *A Nation at Risk*, cited a decline in student scores on SATs from 1963 through 1980 of over 40 points on average in mathematics. The report further documented the poor performance of American students in comparison to other students in industrialized nations, while it also stated that only 33% of students could solve a mathematics problem requiring multiple steps. The report touched off a wave of educational reform that laid the foundation for the development of standards based mathematics curriculum in which performance could be measured using standardized assessments. In the aftermath of the *Nation at Risk* report, the National Council for the Teachers of Mathematics (NCTM) drafted the first set of national mathematics standards for students in grades K through 12. This document became the focus of the "mathematics wars" during the 1990s as proponents, which generally included educational administrators and professors of education, argued that these standards would ensure conceptual understanding in mathematics by utilizing a constructivist approach to explore "real world" problems. Detractors of NCTM standards (including parents and mathematicians) cited the standards' de-emphasis on computational skills noting that the use of calculators was encouraged at all grade levels. Furthermore, opponents criticized the NCTM standards by noting that topics like arithmetic and algebra were not stressed, while other mathematics topics were redundant across grade levels. Despite significant criticism, the NCTM Standards published in 1989 became the basis for the public education mathematics standards developed in most states by 1997. In 2000, the NCTM published a revision of standards for mathematics designed to address the central criticism that the standards did not encompass basic skills. This document was titled *Principles and Standards for School Mathematics*, and it served as the foundation for the mandate of standards based accountability for mathematics under NCLB.

The Impact of No Child Left Behind

Mann and Shakeshaft (2003) stated, "No Child Left Behind has made states responsible for school districts, and school districts responsible for schools" (p. 20), by requiring annual testing of students, academic improvement (measured by adequate yearly progress), and public report cards for all schools. These changes were in stark contrast to previous uses of standardized testing by states and school districts that lacked any real consequences for poor performance, but did require public reporting. Two notable impacts these changes have wrought were the elevation of annual assessment data to an educational commodity of premium value and the ability to analyze this data commodity in aggregated and disaggregated formats for the purposes of making decisions related to school change (Dougherty, 2002). This process of gathering and analyzing data to guide decisions in education has become known as "data-driven decision making" (Marsh, Payne, & Hamilton, 2006, p. 1). In past decades, schools have relied on administrator's intuition and anecdotal information to make decisions about school change. Successful practices were often copied from other school districts under the assumption that they would work in any school without regard to the needs indicated by the study of unique data (Jennings, 1999).

Killon and Bellamy (2000) pointed out that standardized test scores and end of semester report card grades have been primary sources of feedback for students that are often too slow in their turn-around to have any relevance. Students have moved on, the aggregated data yielded little information about practices and factors that produced better scores (Killon & Bellamy, 2000). Technology has altered the landscape of data use through the ready availability of tools that have the potential for storage and continuous analysis of vast amounts of data. In addition, increased requirements by state departments of education and federal legislation like the *No Child Left Behind Act* (2001) have mandated an ever increasing amount of data to be reported by schools. These requirements have forced schools to establish data collection systems, and laid the foundation for the practice of data-driven decision making in schools.

The pressure that school principals perceive under the *No Child Left Behind Act* to achieve performance in mathematics through the use of data is well documented (Davis, 2006; Johnson, 2006). The pressure to demonstrate proficient performance on state standardized assessments is exacerbated by continued poor performance by U.S. secondary aged students on global measures like the Programme for International Student Assessment (PISA) administered in 2006. U.S. students ranked 26th in mathematics proficiency of 30 participating counties. Izumi (2009) pointed out that there is a breakdown in the use of this data on mathematics performance to influence change in schools and in classrooms.

Though mathematics score data receives much attention at local, state, and national levels, standardized test score data is not the only type of data used in schools. Schools collect many types of data in accordance with state requirements in efforts to understand what may or may not influence standardized assessment scores (Streifer, 2004). Research by Bernhardt (2007) indicated that using different types of data in a system of continuous improvement at any school level will lead to increased student achievement in subjects like mathematics. But is there a difference in the use of data systems by secondary principals to increase student achievement in mathematics on standardized assessments in high and low performing schools? What types of data do these principals perceive to be most effective when making decisions related to mathematics instruction? What data tools do these principals use? This study sought to examine the differences in the use of data driven decision making frameworks by secondary principals in high and low performing schools.

Statement of the Problem

The purpose of this study was to compare the use of data systems by secondary principals, and their perceptions of data type and data tool effectiveness in enhancing student achievement in mathematics in high and low performing schools as measured by the percentage of grade 11 students rated proficient on the 2008-2009 PSSA standardized mathematics assessments.

Research Questions

During the course of this study the following inquiries guided this investigation:

1. What types of data do secondary principals in high and low performing schools use and perceive as effective for increasing student mathematics performance on standardized assessments?
2. What types of data tools do secondary principals in high and low performing schools use and perceive as effective to enhance mathematics scores on standardized assessments?
3. To what extent do secondary principals in high and low performing schools use data systems to enhance student achievement in mathematics?
4. How do secondary principals' uses of data driven decision making in high and low performing schools differ?
    a. Is there a difference between secondary principals' use of data types in high and low performing schools?

b. Is there a difference between secondary principals' perception of data type effectiveness for enhancing mathematics scores in high and low performing schools?

c. Is there a difference between secondary principals' use of data tools in high and low performing schools?

d. Is there a difference between secondary principals' perception of data tool effectiveness for enhancing mathematics scores in high and low performing schools?

e. Is there a difference in the extent that secondary principals in high and low performing schools use data systems to enhance student achievement on standardized mathematics assessments?

Significance of the Problem

Accountability requirements under NCLB have inspired schools to implement and use data-driven decision making at an increased pace. The increased use of data by principals to make decisions in their professional practices is supported by research (Armstrong & Anthes, 2001; Englert, Fries, Goodwin, Martin-Glenn, & Michael, 2004). Increased use of data by principals has been largely in response to the call for increasing standardized test scores in mathematics and reading as principals feel the pressure to meet adequate yearly progress under NCLB (Davis, 2006). Poor results by students on national assessments like the PISA and the TIMMS have increased pressure for improvement in secondary mathematics (Hopkins, 2007; Schmidt 1998). It has become clear that the success of principals under the standards based accountability system mandated by NCLB will depend in part on their ability to use data systems (Halverson, Prichett, Grigg, & Thomas, 2005; Halverson, Prichett &Watson, 2007; Meadows, 2008). Several studies found that instructional leadership utilizing data is critical to student achievement (Leithwood, Louis & Anderson, 2004; Halverson, et al., 2007).

Marsh, Pane, and Hamilton (2006) reported that data was used most frequently by principals to inform decisions related to the assessment of client needs or weaknesses like identifying student weaknesses pertaining to state standards or needs for teachers regarding professional development. Though principals may use data to identify skill gaps most frequently, research indicated that middle and secondary mathematics teachers were found to be the least likely core subject teachers to use data systems in this way and showed a statistically significant difference of less use when compared to elementary teachers (Means, Gallagher, & Padilla, 2007). Mathematics teachers were reported to be less likely to have received professional development for the use of data in their classrooms (Means, et al., 2007). Though teacher competency has been identified as having the greatest influence on student achievement (Marzano, 2003; Sanders, 2000), instructional leadership by principals through the use of data systems has been found to be important in student achievement overall (Marzano, Waters, & McNulty, 2005), and specifically in student mathematics achievement on the PSSA assessment (Merlino, Bernotsky, Feldman, & Rui, 2009).

Research on using data to inform decisions by principals has focused on decision making using standardized score data to understand how school processes might be enhanced to improve outcomes like test scores (Foley, Mishook, Thompson, Kubiak, Supovitz, & Rhude-Faust, 2008). However, studies showed that high performing schools use multiple sources of data regularly for decision making to increase student achievement (Alonzo, 2006; Anderson, 2005; Light, Wexler, & Heinze, 2004). Bernhardt (2007) contended that the most effective decisions are made at the points where different data types intersect because the use of multiple data types presents a clearer picture. Several frameworks for the practice of effective data driven decision making in schools encompass the use of input, outcomes, process, and satisfaction data (Bernhardt, 2007; Mandanich, Honey, Light, & 2006; Marsh, et al., 2006). Although studies showed that teachers and principals rely primarily on standardized assessment data over other available data to inform their decision making most often (Davis, 2006; Dembosky, Pane, Barney, & Christina, 2005; Means, Padilla, DeBarger, & Bakia, 2009), supporting the contention that this single type of data is used to inform several types of decisions in schools (Englert, Fries, Goodwin, Martin-Glenn, & Michael, 2004; Hamilton, Berends, & Stecher, 2005).

Several studies have examined student score data on standardized tests (Merlino, et al., 2009; Phillips, Norris, Osmond, & Maynard, 2002), but they do not suggest relationships between success on standardized tests and the use of data systems by principals to improve mathematics scores. Research by Phillips, et al. (2002) and the work of others emphasized a correlation with initial student score data and subsequent student score data (Rescorla & Rosenthal, 2004) or with academic success by another measure like grades (Willingham, 2002). Other related research has explored variables like social and demographic factors related to student score data over time (Bali & Alavarez, 2003; Heck, 2000; Merlino, et al., 2009), teaching effectiveness (Sanders & Horn, 1998) or test construction itself (Cole, 2001).

All public school principals are held to a measure of accountability under NCLB. It is clear that virtually all principals use data and data tools in some measure to inform decisions. Few studies have sought to compare principals' level of engagement in these practices in high and low performing schools. Marsh, et al. (2006) cited such a gap in the research by pointing out that few studies have examined the impact types of data informed decisions have had on standardized assessment scores. Perceptions of their own practices related to data-driven decision making can only be understood by studying principals. Educators in public schools required to participate in standardized testing may find the information gained from such an analysis valuable for its potential impact and meaning for the data-driven decision making process practiced in their school as it relates to student proficiency in mathematics measured by standardized assessments.

Theoretical Framework

NCLB has mandated accountability for schools that is measured by high stakes testing in mathematics and reading. Since the implementation of NCLB, students in Pennsylvania and across the United States have shown little to no improvement in mathematics by state and national assessment measures. The legislation is founded on the notion of accountability through regular assessment coupled with consequences for failure. This external measure has had significant impact on internal accountability systems of schools (Davis, 2006; Kloby, 2007; Johnson, 2006). In order to meet the demands of NCLB accountability, principals have been forced to develop skills in using data (Halverson, et al., 2005; Halverson, et al., 2007).

Principals' aptitude in understanding the interaction of variables within a system (e.g., school building) is linked to the evaluation of their school by standardized test scores. The capability to view the system as a whole, and understand the interaction of the parts is critical to establishing a process of continuous improvement as described by systems theory (Senge, 1990). This understanding is fundamental to successful use of data because the most powerful informed decisions are made based on multiple types of data at points where the data intersect (Bernhardt, 2007). The use of data in this way has become part of high quality school leadership, and a necessary condition for increasing student achievement (Bernhardt 2007; Halverson, et al., 2007). Principal leadership has a significant influence on student achievement (Leithwood, 2004; Marzano, et al., 2005). The use of data by school principals is recommended by NCLB and required by the Pennsylvania Department of Education for struggling schools to inform efforts at increasing student achievement on standardized assessments. It is clear that virtually all principals use data to some degree, however, variation in PSSA assessment results for mathematics indicate that some principals use data differently.

Overview of Methodology

This study used methodology in two parts. In the first part, PSSA exam data for the school year 2008-2009 was analyzed for all Pennsylvania secondary schools to identify the lowest scoring 150 public secondary schools and the highest scoring 150 public secondary schools as measured by the percentage of grade 11 students who scored proficient or above in mathematics on the exam. PSSA exams are evaluated for reliability and validity. Reliability estimates are generated and published by the Pennsylvania Department of Education. For example, mathematics assessments for 2007 had high and uniform reliability estimates as determined by Cronbach's Coefficient Alpha ranging from .923 to .931 in grades 5, 8, and 11 (Pennsylvania Department of Education, 2008).

The PSSA focuses primarily on content validity. The content validity of the assessments was founded in the work of expert judgments in addressing the Pennsylvania State Board of Education's Academic Standards (Pennsylvania Department of Education, 2008). These individuals are generally teachers and curriculum experts from a representative sampling across the state, who construct test items that are submitted for committee review and field tested to determine the content suitability for possible use on a PSSA exam (Pennsylvania Department of Education, 2008). The PSSA was found comparable to instruments used in other studies (Anderson, 2006; Rescorla & Rosenthal, 2004). In the second part, secondary principals in the identified low and high performing schools were asked to participate in a survey that collected data related to the principals' use of data types, data tools, and data systems.

Definition of Terms

Acequate Yearly Progress (AYP)

An individual state's measure of yearly progress toward achieving state academic standards that: (a) applies the same high standards of academic achievement to all public elementary school and secondary school students in the state; (b) is statistically valid and reliable; (c) results in continuous and substantial academic improvement for all students; (d) measures the progress of public elementary schools, secondary schools, and local educational agencies and the state based primarily on the academic assessments described in paragraph (3); (e) includes separate measurable annual objectives for continuous and substantial improvement for each of the following:

(I) The achievement of all public elementary school and secondary school students.

(II) The achievement of---

a. Economically disadvantage students;
b. Students from major racial and ethnic groups
c. Students with disabilities; and
d. Students with limited English proficiency. (NCLBA, 2001, p. 23)

Data-driven decision making—The systematic collection of information pertaining to school, district, and state goals and standards for the purpose of analysis and interpretation of the data to inform decision making (Bernhardt, 1998).

Data tools—Computer based applications designed to organize, store, and analyze various types of school and student information (Bernhardt, 2007).

Data types—Four categories for organizing information in schools that are titled input, process, outcome, and satisfaction. Categorization of data in this way serves to help schools determine if they are meeting student needs by analyzing the interaction of data within these categories (Bernhardt, 2007; Marsh 2006).

DMAs—Data Management and Analysis systems. Software based systems designed to provide means for the access, organization, and analysis of data (Wayman, Stringfield, Hopkins, & Yakimowski, 2004).

Principalship—Essential performances of a school principal (Gross, 2007).

Principal—A building-level administrator in a school district (Gross 2007).

Secondary Principal—A building-level administrator in a school working in a building that includes grade 11.

Proficiency—Term used to describe students that have mastered Pennsylvania's assessment anchor content standards at their grade level as demonstrated by their score on the PSSA exam(s) (Pennsylvania Department of Education, 2003).

PSSA—Pennsylvania System of School Assessment is a standards-based assessment administered in all public schools for grades 3-8 and grade 11 in the curriculum areas of mathematics, reading, writing and science (Pennsylvania Department of Education, 2003).

No Child Left Behind Act of 2001—Federal law that requires states to adopt a program of high-quality, annual student assessments that can be used to determine the annual performance of every school and local education agency in exchange for continued federal funding (NCLBA, 2001).

PSSA Exam—Another word for *test*. Under *No Child Left Behind*, tests are aligned with academic standards. Schools must administer tests in each of three grade spans: grades 3-5, grades 6-9, and grades 10-12 in all schools. Since the start of the 2005-2006 school year, tests have been administered every year in grades 3 through 8 and in grade 11 in mathematics and reading (Pennsylvania Department of Education, 2003).

Primary Grades—School grades kindergarten through grade eight.

Secondary Grades—School grades 9 through grade 12.

Standardized Test—A way of giving, scoring, or reading tests so that the data taken from all of them can be compared (Pennsylvania Department of Education, 2003).

Value-Added Models (VAMs)—Models that employ statistical analysis to interpret standardized test scores and other factors related to student achievement to determine and project academic growth over time (Rubin, Stuart, & Zanutto, 2004).

Summary

NCLB has significantly impacted school principal perceptions regarding accountability (Davis, 2006; Johnson, 2006). The use of standardized test score data as a measuring stick under NCLB has marked a change in the required skill sets of school principals. In this era of high stakes testing, principals must be fluent in the use of data systems to make decisions in their schools in order to be successful under the current accountability measures (Halverson, et al., 2007).

Standardized test scores of students in the state of Pennsylvania are a microcosm of national assessments. On average the percentage of students scoring proficient on the PSSA mathematics assessment has declined from the primary to secondary grades. While mathematics scores on this assessment for grade 11 students have marginally improved, they still lag behind student performance in reading. The use of data systems to make decisions in schools has been cited as the most appropriate way to address problems in schools including the issue of raising student achievement (Bernhardt, 2007; Schmoker, 2006; Streifer, 2004). Research shows that the use of these systems by principals has significant effect on student achievement (Marzano, et al., 2005).

CHAPTER 2

REVIEW OF RELATED LITERATURE

NCLB was designed to ensure that all students have an equal opportunity to obtain high quality education that provides them the necessary skills to reach at least a proficient level of performance on challenging state academic assessments. Required yearly assessments in mathematics, reading, and science coupled with consequences for poor performance have significantly impacted perceptions of accountability in schools as school principals feel the pressure to reach required levels of proficiency on state assessments (Davis, 2006). Efforts to meet the demands of accountability under NCLB have required school principals to develop new skill sets that include establishing their own proficiency in the collection and analysis of standardized test score data (Bernhardt, 2007; Englert, et al., 2004). The development and implementation of this data skill set has become a high priority for many secondary principals in states like Pennsylvania where student proficiency levels in mathematics have lagged significantly behind proficiency levels in reading since the passage of NCLB (Pennsylvania Department of Education, 2009). There is an even greater disparity between Pennsylvania secondary students' levels of proficiency in mathematics when compared to proficiency levels of students in grades three through eight that reached nearly 16% last year (Pennsylvania Department of Education, 2009). Performance levels by Pennsylvania secondary students are a microcosm of global performance by U.S. students in mathematics. American students have demonstrated superior performance over many countries in reading, while their performance in mathematics and science ranks much closer to the bottom in comparison to other industrialized nations (Hoff, 2001; Hopkins, 2007).

Heritage and Chen (2005) argued that improvement of student mathematics scores on standardized assessments rests largely on the effective use of data in schools, and that principals will have a significant role in this process. This contention was supported by research regarding the use of data in schools to support efforts designed to increase student performance in mathematics. Numerous studies have examined supporting factors and practices of data-driven decision making in schools (Marsh et al., 2006; Means, et al., 2009; U.S. Department of Education, 2008), but few have compared the level of principal engagement in this process and achievement on standardized mathematics tests in high and low performing secondary schools. To provide a framework to address these areas of further study, relevant research on data-driven decision making, systems theory, and NCLB were reviewed. In total, the interrelation of Senge's (1990) systems theory, the process of data-driven decision making, and mandates set forth under NCLB, served to describe the process of mandated standardized testing, the foundations of data-driven decision making process, and what may be gained from comparison of secondary school principals' level of engagement in this process in high and low performing schools as measured by standardized mathematics assessments.

No Child Left Behind Act

The *No Child Left Behind Act of 2001* is rooted in the theory that schools must be held accountable for student performance. Accountability is not new to schools. However, this particular Act has made the "*new accountability"* as described by Fuhrman (1999) into law. This new accountability has common components including: (a) emphasis on student outcomes as the measure of adult and system performance; (b) a focus on the school as a basic unit of accountability; (c) public reporting of student achievement; and, (d) the attachment of consequences to performance levels (Abelmann, Elmore, Even, Kenyon, & Marshall, 1999;

Fuhrman, 1999). These are all components found within NCLB that support the theory that holding schools and districts accountable according to this model, determined by NCLB, will lead to school improvement.

NCLB has provided for established systems of standardized assessments designed to provide an avenue for school district and individual school accountability including requirements that force school change based on school progress toward NCLB goals as indicated by assessment data (No Child Left Behind Act, 2001). The NCLB Act established five major accountability requirements for states: (a) assessment of students in grades 3-8, and once during grades 10-12, in reading and mathematics by the end of the 2005-2006 school year, using state designed tests; (b) certification that all teachers of core academic subjects are highly qualified by the end of the 2005-2006 school year, using state-set definitions; (c) assessment of students in science once during grades 3-5, 6-9, and 10-12 by the end of the 2007-2008 school year, using state designed tests; and, (d) provision of public school choice and supplemental educational services to students in schools that have been unable to meet Adequate Yearly Progress for two consecutive years; accept nothing short of 100% student proficiency by 2014 (No Child Left Behing Act, 2001).

Internal and External Accountability

Prior to the implementation of NCLB, Abelmann, et al., (1999), conducted a study of 20 schools of mixed variety including urban, rural, private, public, and charter, and examined how schools construct individual conceptions of accountability as part of the school's identity. The study was partially based on work by Newmann, King, and Rigdon (1997), which suggested that agencies seeking to establish accountability for education should pursue stimulation of internal accountability versus the assumption that accountability can be established by exercising external measures alone. Abelmann, et al. found that external influences (e.g., laws and curriculum mandates made by central offices) are only one out of many factors that influence a school's internal concept of accountability. Though many schools selected for this study existed in some type of formalized accountability structure prior to the passage of NCLB, the research indicated that these structures had no effective influence on teachers' knowledge of for what and to whom they were accountable.

Abelmann, et al. (1999) identified three factors, which interacted to comprise a school's internal accountability system. They were individual responsibility, collective expectations, and accountability. *Individual responsibility* was defined as what teachers perceived as their degree of responsibility for student learning. *Collective expectations* were expectations about the behavior of others that exceed individual responsibility, and *accountability* was the, "formal and informal ways in which people give an account of their actions to someone in formal authority, in or outside of the school" (Abelmann, et al., p.13). Five schools noted in the study evidenced strong internal accountability structures that were both formal and informal in nature, and had clear collective expectations for all involved. However, other schools evidenced weak or non-existent accountability structures in which teachers essentially believed that they were accountable only to themselves. In other words, the accountability structure in these schools was defaulted to individual responsibility in the absence of collective expectations. As a result, these researchers found internal accountability structures in these schools fragmented or non-existent. Regardless of the accountability structure that developed in a school, the study asserted that schools, "develop their own internal normative structures that are relatively immune to external influences" (p. 48). Abelmann, et al., (1999) concluded that the individual school as the defined unit of accountability would succeed or fail depending on how closely the internal accountability structure that developed is matched with the requirements of the external accountability structure.

Post NCLB research indicated a change in the influence of external accountability measures on the internal accountability structures in schools because accountability is now driven by high-stakes testing as defined under NCLB (David, 2006; Hamilton, et al., 2005; Johnson, 2006; Kloby, 2007; LoGerfo, 2004). Davis' study included 310 of 514 middle and high school principals working in public schools in Tennessee. His research showed that all principals perceived and reported that more time was now devoted to teaching of subjects that are part of high stakes testing as mandated by NCLB than subjects that are not tested. Furthermore, principals perceived that this increased time may show results in the form of higher standardized test scores, but they did not believe it improved the quality of learning or was an accurate measure of their school's effectiveness.

LoGerfo's study (2004) of 9,744 first-grade students and 2,390 first-grade teachers in 697 schools found a strong correlation between teachers with a higher sense of responsibility for their students' achievement and increased student performance. LoGerfo asserted that teachers who had a high degree of responsibility with regard to their students' learning, have internalized the external measures of accountability found in NCLB as part of their personal responsibility for their students' learning.

A study of teacher and principal responses to standards-based accountability by Hamilton et al. (2005) further attested to the influence of NCLB on classroom practice. They surveyed 6,672 teachers at the elementary and middle levels in three states including 2,355 from the state of Pennsylvania. This study showed that 76% of the Pennsylvania teachers responded that their instruction was affected by the state mathematics assessment established and administered as required by NCLB. Furthermore, 28% of the Pennsylvania teacher sample also reported that they spent more time focusing on students who were close to proficiency, while 75% reported that they had increased focus on topics emphasized on the assessment. Clearly, NCLB though an external influence, has influenced school internal accountability systems.

Standards Based Accountability and Mathematics

A National Research Council document described common components of the standards based accountability system as established under NCLB:

> The centerpiece of the system is a set of challenging standards. By setting these standards for all students, states would hold high expectations for performance; these expectations would be the same regardless of students' backgrounds or where they attended school. Aligned assessments to the standards would allow students, parents, and teachers to monitor student performance against the standards. Providing flexibility to schools would permit them to make the instructional and structural changes needed for their students to reach the standards. And holding schools accountable for meeting the standards would create incentives to redesign instruction toward the standards and provide appropriate assistance to schools that need extra help. (National Research Council, 1999, p. 2-3)

Research indicated that the shift to a standards based accountability system had considerable impact on the practices of mathematics teachers and school principals (Hamilton, Stecher, Marsh, McCombs, Robyn, Russell, Naftel, & Barney, 2007; Hamilton, 2004). Principals reported an increased use of data to direct school improvement for subjects like mathematics, which they attributed to standards based accountability requirements under NCLB (Hamilton, et al., 2007). After conducting a multiyear study in three states including Pennsylvania, Hamilton, et al. (2007) found that mathematics teachers reported a significant increase in the use of teaching strategies related to the individualization of instruction and the use of data. Nearly 20% of mathematics teachers reported increasing their use of the following strategies: providing individual help to students outside of class time; conferring with another teacher about ways to present specific topics or lessons; having students help other students learn the content; planning different assignments or lessons based on student performance; reteaching mathematics topics because student performance on assessments or assignments did not meet expectations; and, reviewing assessment results to identify individual students who need supplemental instruction. Additionally, over 50% of mathematics teachers in this study reported that they used data gathered from state assessments to be helpful in the following areas: identifying teacher content knowledge deficiencies, identifying gaps in curriculum and instruction, and tailoring instruction to individual student needs. These findings were supported in a national study by the United States Department of Education (2008) that demonstrated a significant increase in the use of data by mathematics teachers in these areas between 2005 and 2007. The use of data in this way has been shown to improve mathematics scores on standardized assessments (Bernhardt 2007).

Results from the National Assessment of Education Progress (NAEP) given to a sampling of students across the nation in grades 4, 8, and 12 indicated that the standards based accountability system under NCLB is producing the intended results for some students in mathematics (Glod, 2009). In 2008, mathematics scores for students in grade 4 and 8 were the highest they have ever been, while scores for high school aged students remained flat at the same level they have been since 1973. While the NAEP assessment is not given to 12th grade students in Pennsylvania, results from the grade 11 PSSA assessment in mathematics over the past eight years have demonstrated nearly flat line results with a slight increase from 49% to 56% of students scoring proficient in mathematics (Merlino, et al., 2009). Nationally, on the 2009 NAEP mathematics assessment, the upward trend for 8^{th} grade students in mathematics continued, while grade 4 and grade 12 mathematics scores remained unchanged. Pennsylvania students in grades four and eight have shown consistent improvements in mathematics in over past six administrations of the NAEP assessment, while also scoring above the national average score.

New Leadership for No Child Left Behind

The role of the principal has changed drastically in the past decades (Farkas, Johnson, Duffett, Foleno, & Foley, 2001; Murphy, 2002). Leaders must still have a deep knowledge of issues such as facilities, personnel, and finance management. However, these areas alone are not enough to support deep, sustained school improvement. Senge (1990) described the traditional view of leadership as being "deeply rooted in an individualistic and nonsystemic worldview" (p. 315) in which leaders serve as heroes in all situations. In a school setting, this concept plays out as principals serving as white knights who exist to fix all problems single handedly.

Senge noted that the traditional view of leadership distorts the focus of organizations preventing them from developing into learning organizations in which members are responsible for learning. There are three key capacities noted by Senge that leaders must master in what he describes as a new leadership for building learning organizations that include: leader as designer, leader as steward, and leader as teacher. Senge explained that the most successful leaders do not fulfill the role of hero; instead they focus their efforts on designing the learning process so that members of an organization can deal with the issues they face, and experience personal growth rather than continually having the leader solve their problems. Stewardship requires that leaders have a clear understanding of the larger picture within an organization including the systems that comprise it so they continually develop the purpose and direction of the organization. Leaders as *teacher* is fostering learning for all, not simply teaching others the vision for the organization.

Fullan (2007) characterized Senge's view of leadership in the changed role of the modern school principal by stating:

> The old world is still around, with expectations that the principal will run a smooth school and be responsive to all; simultaneously, the new world rains down on schools . . . expecting that at the end of the day the school constantly should be showing better test results . . . . (p. 157)

Fullan suggested that this "new world" has resulted in the role of school principal to be constantly increasing in complexity, while driving many educators away from pursuing the position. Part of what Fullan notes as the "the new world rains down" are the requirements of high stakes testing under NCLB. Though demands of the principalship are rigorous and many, Fullan pointed out that a common factor in successful schools and in schools showing improvement is strong principal leadership.

Research supports Fullan's contention that success in schools is connected to effective leadership (Duke, 2004; Leithwood, et al., 2004; Marzano, et al., 2005). In a review of literature for their study on the influence of principal leadership on student learning, Leithwood, et al. (2004) described three core practices by successful school leaders that emerged. These core practices mirror those noted by Senge (1990) and were described as: setting directions (leader as *steward*); developing people (leader as *teacher*); and, redesigning the organization (leader as *designer*). The study of 2,570 teachers conducted by Leithwood, et al. in 90 elementary and secondary schools demonstrated the importance of school leadership in relationship to student learning. Results from the study showed that one quarter of the variation of student achievement explained by school level variables can be attributed to school leadership effectiveness. Based on their review of 69 studies, Marzano, et al. (2005) supported this correlation between effective school leadership and student achievement.

The importance of effective principal leadership has been recognized and acknowledged by various educational entities including the Council of Chief State School Officers (CCSSO). In 1996, the CCSSO established the Interstate School Leaders Licensure Consortium (ISLLC) that drafted and adopted standards for school leaders. These standards were designed to establish "a common core of knowledge, dispositions, and performances" (p. iii) that will facilitate a stronger connection between leadership and successful schools.

The ISLLC Standards are as follows:

Standard 1. A school administrator is an educational leader who promotes the success of all students by facilitating the development, articulation, implementation, and stewardship of a vision of learning that is shared and supported by the school community.

Standard 2. A school administrator is an educational leader who promotes the success of all students by advocating, nurturing, and sustaining school culture and instructional program conducive to student learning and staff professional growth.

Standard 3. A school administrator is an educational leader who promotes the success of all students by ensuring management of the organization, operations, and resources for a safe, efficient, and effective learning environment.

Standard 4. A school administrator is an educational leader who promotes the success of all students by collaborating with families and community members, responding to diverse community interests and needs, and mobilizing community resources.

Standard 5. A school administrator is an educational leader who promotes the success of all students by acting with integrity, fairness, and in an ethical manner.

Standard 6. A school administrator is an educational leader who promotes the success of all students by understanding, responding to, and influencing the larger political, social, economic, legal, and cultural context.

The use of data in schools has become an integral part of the new principal leadership as defined under the ISLLC standards. The requisite knowledge, disposition, and performance pertaining to the use of data by school leaders is well defined under Standard 1 by stating under the subheading of Knowledge that school leaders should "have knowledge and understanding of information sources, data collection, and data analysis strategies" (p. 10). Contained under the subheading of Dispositions for Standard 1, it is noted that school leaders should "believe in, value, and be committed to continuous improvement" (p. 10). Performance indicators for Standard 1 indicate that the use of data in goal and mission setting in schools is crucial including the use of student assessment and demographic data. Research by Gross (2007) suggested that Pennsylvania principals view the ISLLC standards as important. Additionally, Gross found that these principals believe their respective preparation programs have adequately trained them to fulfill expectations outlined by the ISLLC standards.

Requirements under NCLB have served to increase the complexity of the principal's role as described by Fullan (2007) by placing a premium on the use of data in schools. Wilson (2004) noted that successful accountability systems in schools master and establish the ability to connect classroom practices and external accountability measures by providing for information or data flow between them. Regular classroom practices must be tied to school wide outcomes, while data systems that provide local and continuous measures of student learning to teachers must be developed in schools (Halverson, et al., 2007). The ability of school leaders to facilitate a change in culture from a long existing internal accountability system to the development of new systems that meet the demands of the external accountability system is at the center of a new instructional leadership emerging in schools (Halverson, et al., 2005).

Halverson, et al. (2005), defined the “new instructional leadership” of the principal by stating:

> The new instructional leadership will require knowledge and frameworks to guide leaders in creating schools that systematically improve student learning. Leaders will need to be able (a) to work with teachers to help students test well, while not reducing learning to testing, and (b) to justify changes in instructional and personnel practices to an increasingly well informed community. (p. 5)

To develop a basis for understanding how requirements for this new instructional leadership might be met, Halverson, et al., studied the practices that principals in schools recognized for improving test scores used to develop *data-driven instructional systems* (DDIS) comprised of six functions including: data acquisition; data reflection; program alignment; program design; formative feedback; and, test preparation. These functions constitute a DDIS framework that serves to describe how leaders create systems that change school culture to meet demands of NCLB accountability. Research by Halverson, et al., demonstrated that effective principal leadership in this new era of accountability can be measured by determining how successful these functions operate within a school as applied to various subjects. Halverson, et al. found that each school in the study demonstrated use of the six functions noted, which resulted in highly developed formative assessment models for language arts. However, the study showed that these schools were considering achievement data in mathematics, but there was little evidence of data use in the same way to address the complexities of mathematics teaching or learning. The finding by Halverson, et al. that effective instructional leadership for principals under NCLB requires effective use of data systems to improve student outcomes is supported by other studies (Meadows, 2002; Torrence, 2002).

Teacher Certification under No Child Left Behind

While there is a connection between principal leadership and student performance, there is also a link between teacher certification and teacher effectiveness measured by student achievement. This finding supports NCLB requiring teacher certification in states so that all core subject area teachers and paraprofessionals are highly qualified under the law. This means that these teachers must acquire state certification through an approved program. Nyankori (2005) found that the elementary students of teachers holding state-issued certification that satisfied the requirement under NCLB had increased gains in mathematics and reading when compared to the achievement of students who had non-certified or emergency certified teachers.

Similarly, research by Vandervoort, Amrein-Beardsley, and Berliner (2004) showed that students assigned to Nationally Board Certified teachers demonstrated increased achievement as measured on standardized assessments when compared to students assigned to teachers with standard certification. Further research by Marzano (2003) identified 19 school level factors that influence student learning. He stated that teacher effectiveness has more impact on student achievement than any single factor. NCLB has further influenced school accountability by requiring more than performance on standardized assessments.

Longitudinal Data Systems

NCLB also requires the standardized testing of students conducted across primary and secondary grades. Though there is little research on such data (Rescorla & Rosenthal, 2004), this particular requirement has initiated interest in collection and analysis of longitudinal data on student performance across school grades. The law does not require states to have specific database structures, but it endorses databases that link students' test scores, the length of time they have been enrolled in given schools and graduation records over time (Education Commission of the States, 2002). The Education Commission of the States (ECS) cited that longitudinal analysis of student performance data at the student-unit level provides the best, most accurate information for both policy decisions and decisions at the district and school levels. The ECS also noted that this analysis is superior because it can provide information about student growth over time that can be tied to teachers, schools, and curricula that served those students. It can also provide fairer comparisons among schools because it ensures school performance is based only on students who have continuously enrolled in that school, rather than comparing cohorts that do not account well for dropouts and transfer students.

In 2005, the National Center for Educational Accountability launched the Data Quality Campaign. Inspired by NCLB, it is a national collaborative effort designed to provide resources and tools for states to develop longitudinal data systems that improve the quality, collection, and use of data to improve student achievement. The National Center for Educational Accountability (2007) identified 10 essential elements of a complete longitudinal data system that all states are encouraged to develop including:

1. A unique statewide student identifier so states may track students over time even if they move from district to district.

2. Student-level enrollment, demographic, and program participation information in order to help measure which programs are helping students succeed, while accurately accounting for transfer students.
3. The ability to match individual students' test records from year to year to measure academic growth.
4. Information on untested students in order to account for all students including those not tested or exempted from testing.
5. A teacher identifier system with the ability to match teachers to students, which is a critical piece in understanding the relationship between teacher qualification and experience and student academic growth.
6. Student-level transcript information that will yield data on courses and grades so that states will be able to track course request patterns and their relationship to success on states, assessments and readiness for post-secondary education.
7. Student-level college readiness test scores that can be stored as an additional data component to provide solid indicators of college readiness and preparedness for the world of work.
8. Student-level graduation and dropout data (currently required data collection in nearly all states presently).
9. The ability to match student records between Pre K-12 and higher education systems will provide secondary schools the ability to connect student performance in high school to student performance in college resulting in a better aligned curriculum that supports increased student preparedness for post-secondary education and work.

10. A state data audit system should be in place to ensure data quality, validity, and reliability. (p. 5-6)

State education agencies, including the Pennsylvania Department of Education, have recognized and have taken action in the establishment of longitudinal data systems as endorsed by NCLB (National Center for Educational Equality, 2007).

In 2007, the Pennsylvania Department of Education launched programs establishing a data system that meets the 10 criteria established by the Data Quality Campaign. The Pennsylvania Information Management System (PIMS) stores student demographic data collected and verified by school districts. The PIMS system promised to:

> Establish enterprise-wide systems and processes to streamline data management and utilization; provide multi-year longitudinal data to help teachers and administrators address individual student needs; empowers teachers with state-of-the-art data analysis tools to improve individual student achievement. (Pennsylvania Department of Education, 2007)

According to the Data Quality Campaign's survey in 2009, Pennsylvania reported implementation of data systems that included 8 of the 10 essential elements compared to the reporting of only 2 elements in 2005. In similarity to most other states, Pennsylvania lacks implementation of systems related to elements six and seven from the Data Quality Campaign list. The ability to match individual students' test records from year to year to measure academic growth can now be calculated by data systems in all states.

Adequate Yearly Progress

Whereas there are benefits to longitudinal data analysis, schools, and school districts are held accountable annually by the AYP measurement established at the state level (Anderson, 2005; No Child Left Behind Act, 2001). Schools that do not make their AYP requirement for all students as well as for student subgroups are required to undergo a series of steps to correct their deficiencies. Schools that do not make AYP for two consecutive years are identified as *in need of improvement*. These schools are required to develop a corrective plan as well as offer all students the option of transferring to another school within the district that is not so identified. Pennsylvania schools that do not make AYP must complete and submit for approval standard forms that are aligned with the states "Getting Results!" framework that is designed to facilitate continuous improvement. These forms require that strategies designed to address AYP deficiencies and their implementation be documented and submitted to the Pennsylvania Department of Education for approval.

If a school does not make AYP for a third consecutive year, the school must continue to provide school choice and assistive services like after school tutoring to low achieving students. Schools that do not meet Adequate Yearly Progress for a fourth consecutive year, must continue to offer school choice, supplemental assistive services, and must implement corrective actions to improve the school based on assessment data like replacing certain staff or fully implementing a new curriculum.

If AYP is not made for a fifth consecutive year, a school is identified for restructuring and would be subject to implementing significant alternative governance actions, state takeover, hiring of a private management contractor, converting to a charter school, or significant staff restructuring. Schools, again, are required to continue offering school choice, and supplemental services. Although used as a performance measuring stick for schools, AYP provides only a snapshot of a given cohort group of students at one point in time. Though it does not provide for analysis of student performance for a particular cohort group from primary to secondary grades, AYP as a measuring stick has been a motivating factor in the gathering of data by schools to make informed decisions (No Child Left Behind Act, 2001).

One-Hundred Percent Proficiency

The 100% proficiency requirement of NCLB proffers the theory that every student can reach a specified level of performance. It holds schools accountable for every student by requiring schools meet AYP targets within student subgroups. Students with disabilities, ethnic groups, English as a second language students, and economically disadvantaged students comprise the subgroups defined under NCLB. Numerous studies have indicated that students in some identified subgroups like special education, minorities, or impoverished students traditionally do not score well on standardized tests (Bali & Alavarez, 2003; Ramanathan, 2007). This NCLB requirement does not allow exceptions on any basis, although states are permitted to develop alternative assessments and provide for appropriate assistance on standardized assessments to those students with special needs who have a diagnosed learning or physical disability and require assistance. Consequently, schools and school districts are required to review student performance data by disaggregated subgroups which have made comparison data between these subgroups valuable in attempting to ensure all subgroups meet AYP.

Although NCLB does not provide direction to schools in terms of what decisions should be made by school administrators based on data results or specific database structures that should be used, it does make it clear that data will be used to correct schools identified as needing improvement. Under Section 1116, *Academic Assessment and Local Education Agency and School Improvement*, the NCLB Act states that schools in need of improvement shall receive specific assistance to include assistance in data analysis of student assessments to identify and address problems of instruction. Section 2113 (*State Use of Funds*) of the NCLB Act, reads that grant funds obtained under Section 2111 of NCLB must be used within a number of areas including:

> Encouraging and supporting the training of teachers and administrators to effectively integrate technology into curricula and instruction, including training to improve the ability to collect, manage, and analyze data to improve teaching, decision making, school improvement efforts, and accountability. (2001)

NCLB encourages schools and school districts in its wording and accountability requirements to develop systems to collect, manage, and analyze data to facilitate the process of data-driven decision making.

Value-Added Accountability and No Child Left Behind

In contrast to value added modeling systems, current NCLB requirements mandate comparison using a cross-sectional approach by measuring a performance at a single time. The results are then compared in aggregated and disaggregated formats against the established proficiency requirement for that school year to determine whether a school made adequate yearly progress. Performance is reflected by the overall group as well as identified subgroups by ethnicity, economic status, English language proficiency, and individualized education plan status. Research suggested that this results in measuring social differences in successive groups of students instead of school effectiveness because some schools that are predominantly white, English speaking with little poverty do not have the required number of students in some subgroups to measure proficiency under the Act (Buchanan, 2004; Kim & Sunderman, 2004).

Many non-educational factors like race, gender, or socioeconomic status have been shown to influence student achievement (Bali & Alavarez, 2003; Heck, 2000). Research has revealed that socioeconomic factors can be isolated from influencing student learning using a method known as *blocking* (Hershberg, et al., 2004; Sanders, 2000). In the Sanders' model students serve as their own control over time to determine their progress, thus isolating any influence of non-educational factors. After these factors have been removed, the resulting measure produced is considered to be the student's true achievement growth. This value-added score can be interpreted as a quantitative measure of the educational factors that directly influence student achievement like the student's teacher or school (McCaffrey, Lockwood, Koretz, & Hamiton, 2003; Hershberg, et al., 2004). This approach effectively eliminates the achievement gaps between socioeconomic groups. Accountability for these subgroups as a separate measure is nullified in contrast to current requirements under NCLB.

Many researchers viewed value-added approaches as being a more fair approach to accountability (Sanders, 2000; Sanders & Horn, 1994; Kelly & Monczunski, 2007). Many high performing schools making AYP have been shown to be failing when measured under a valued added approach, whereas many low performing schools as measured by AYP would be considered successful under value-added accountability (North Carolina Department of Public Instruction, 2006; Tennessee Department of Education, 2006). Hershburg (2004) argued that this disparity results in unfair comparisons of effectiveness between schools that could be avoided by using a value-added approach. Marsh, et al. (2006), contended that the AYP mandate has resulted in unfair classroom practices by teachers who have ignored the lowest achieving students in favor of focusing instruction on those closest to reaching a proficient score on the standardized assessments. The lowest performing students, who need the most focused instruction, do not receive it because their gains in learning, though they may be exceptional, will not reach a level to increase the number of students deemed proficient as measured under AYP. Drury and Doran (2003) summarized this point by stating:

> It makes little sense to continue to define AYP solely in terms of the percentage of students crossing an arbitrary bar of "proficiency", while ignoring the growth that occurs within broad performance categories. This is tantamount to measuring a child's height with a yardstick but acknowledging growth only when his or her height exceeds 36 inches. (p. 3)

The longitudinal approach utilized in value-added accountability systems avoids cohort to cohort comparison. The analysis used to determine AYP is often deemed unfair because it takes a snapshot of achievement of completely different sets of students while doing little to explain any increase or decrease in student achievement (Drury & Doran, 2003; Meyer, 1996).

Though there is no provision in NCLB, or in any other law, for the establishment and use of value-added systems for accountability, the United States Department of Education supports the value-added approach. In November 2005 the Department awarded grants to 10 states, including Pennsylvania, for developing value-added models. Pennsylvania used its grant to develop the PVAAS system. All of these value-added models were based on the EVAAS system, which is an extension of the Sanders' model. Tennessee requested, but was not permitted to use the TVAAS system for accountability under NCLB. All 10 states will still be required under NCLB to show all students proficient by 2014, but are permitted to incorporate the value added systems into state accountability. Pennsylvania established 2006 as the base year in which it began providing all schools report data from the PVAAS system, and the system was made available for use by all educators in 2009.

Data-Driven Decision Making

The concept of data-driven decision making is not new to schools, and has been in practice in some measure, since the 1970s and 1980s (Popham, 1987; Popham, 2001; Teddlie &

Reynolds, 2000). Data-driven decision making was founded on taking a systems theory approach to understanding how schools work. Bernhardt (1998) described data-driven decision making in schools as learning continuously and applying what is learned to improve results on a continuous basis. The United States Department of Education (2008) provided an education specific definition:

> Data informed decision making is the analysis and use of student data and information concerning education resources and processes to inform planning, resource allocation, student placement, and curriculum and instruction. The practice entails regular data collection and ongoing implementation of an improvement process. (p. 2)

Bernhardt (2000) further explained the process by noting that the practice of collecting and interpreting data can identify root causes of problems instead of symptoms of problems; assess the resources required to address them; and set goals and measure progress toward them. Numerous studies have indicated that the use of data in this way can lead to school improvement (Alonzo, 2006; Armstrong & Anthes, 2001; Barry, 2006).

School Culture and Data Use

Bernhardt (2007) contended that many schools must undergo a shift in school culture before they can become effective in using data to make decisions because the process requires commitment to continuous improvement. Commitment to continuous improvement using data implies continuous commitment to learning within an organization. Senge (1990) noted the importance of a shared vision within an organization as being a critical component for a learning organization. Senge defined shared vision at its simplest level as, “the answer to the question, What do we want to create?” (1990, p. 206). Bernhardt (2007) stated that effective data-driven decision making must be based on a culture that is committed to continuous improvement which is for most schools a restructuring initiative that requires visionary leadership and an ability to see the whole picture. Bernhardt succinctly noted, “the first step is to create and lead shared vision for the organization that is based on shared values and beliefs and on the purpose of the school and district” (p. 104). This precursor to an effective data-driven decision making process will develop as consistent and fundamental understandings regarding the use of data in supporting the implementation of the shared vision become commonly held within the organizational culture. Research has supported Bernhardt’s contention that a strong shared vision is fundamental to implementing continuous improvement incorporating data in an organization (Alonzo, 2006; Barry, 2006; Mandinach, et al., 2006).

In their case study of six school districts in demographically different systems in three states, Datnow, Park, and Wohlstetter (2007) explored development of school culture as a precondition for effective data-driven decision making. The study included observations, interviews of teachers, principals, superintendents, and collection of standardized score data. The research showed high performing districts were highly engaged in using data to inform decisions, and that significant efforts were made to improve school culture related to the process prior to implementation. These efforts included the creation of explicit expectations throughout the school systems by central office administrators and schools by principals. By modeling these expectations, principals were able to establish perceived accountability for themselves, teachers, and central office administrators. Principals reported that a change in culture was the foundation of their success in using data.

Data-Driven Decision Frameworks

Several frameworks that established models for the data-driven decision making process have emerged from the study of high and low performing schools (Bernhardt, 2007; Breiter &

Light, 2006; Mandinach, et al., 2006; Marsh, et al., 2006). Bernhardt (1998) established a conceptual framework for data-driven decision making that begins with seven questions designed to examine the use of data in schools starting with one question that supersedes all others, “What is the purpose of the school?” (p. 9). Bernhardt asserted this question should be a starting point for schools establishing a culture of continuous improvement. Bernhardt (1998) contended that the remaining six questions are influenced by the answer to the first.

1. What do you expect students to know and be able to do by the time they leave the school? (Standards)
2. What do you expect students to know and be able to do by the end of each year? (Benchmarks)
3. How well will students be able to do what they want to do with the knowledge and skills they acquire by the time they leave school? (Performance)
4. Do you know why you are getting the results you are getting? Do you know why you are not getting the results you want?
5. What would your school and educational processes look like if your school were achieving its purpose, goals, and expectations for student learning?
6. How do you want to use the data you gather? (p. 9)

Bernhardt suggested these questions be considered while evaluating data through “lenses” of demographics, student learning, school process, and perceptions. Senge (2000) emphasized using system maps, which is fundamental to systems thinking, to aid understanding the key elements of systems and how they interact with each other. These diagrams also serve to provide a conceptual framework of variables interacting across various structures and levels within an organization. Bernhardt's (2007) established framework that illustrates the interaction of data within a school organization is shown in Figure 1.

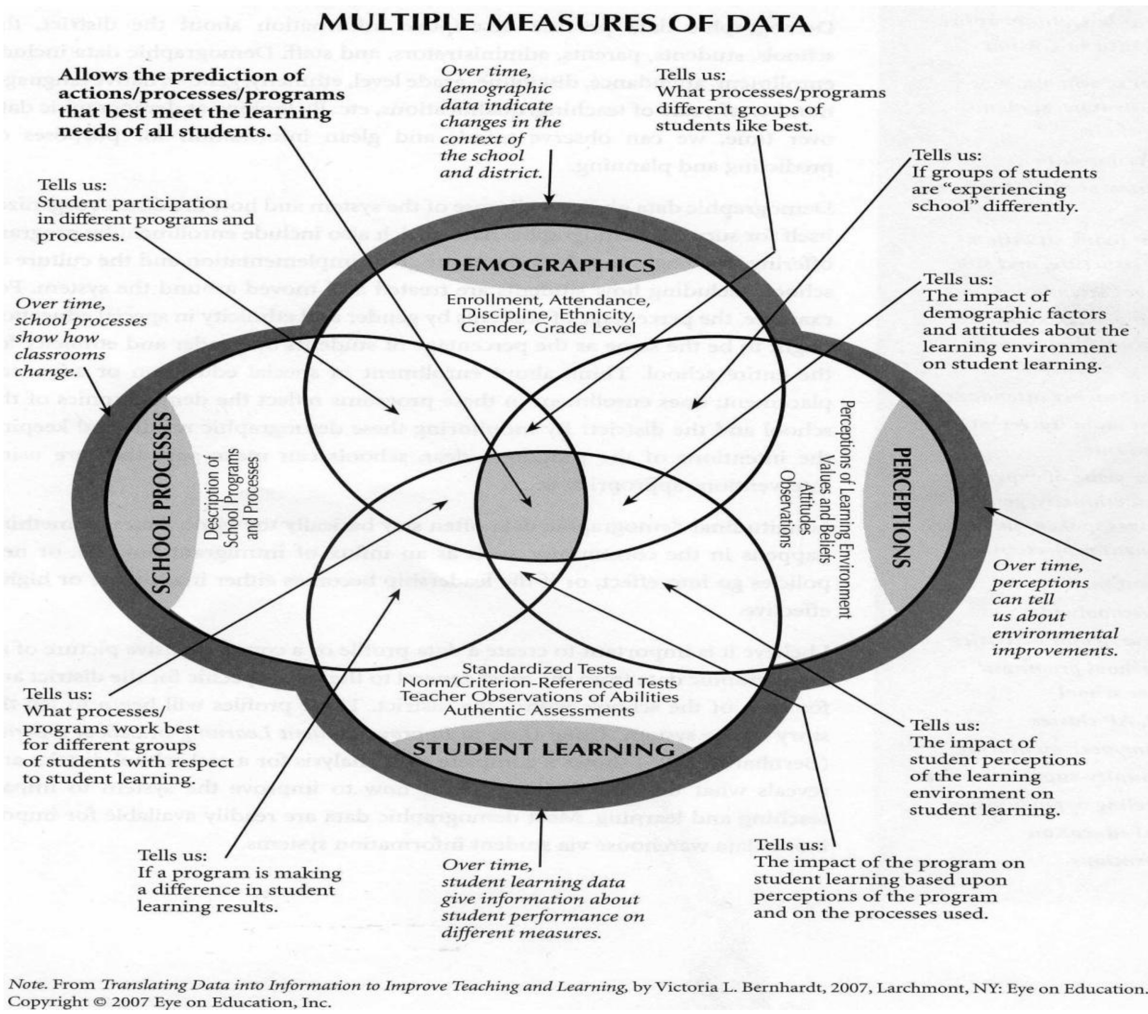

Figure 1. Multiple measures of data: Categorization of data into specific types and illustrates how each category interacts with the others. Informed decisions are made from data gained from the intersection of two or more categories. (From Translating Data into Information to Improve Teaching and Learning by Victoria L. Bernhardt, 2007, Larchmont, New York Eye on Education. Copyright © 2007. Reprinted with permission.)

Bernhardt's (2007) illustration is most powerful when data across the four categories intersect at the center of the framework and schools make decisions based on this data.

Mandinach, et al. (2006) developed the conceptual framework shown in Figure 2 by using systems thinking as an analytical perspective.

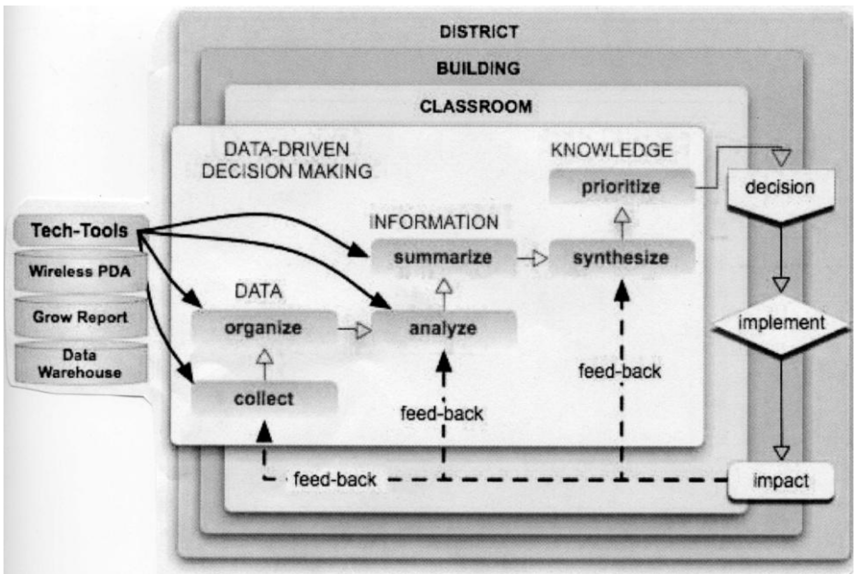

Figure 2. Data informed decision making. Data, information, and knowledge interact to form a foundation for decision making. Feedback further modifies data, information, and knowledge to fine-tune the decisions. (From *A Theoretical Framework for Data-Driven Decision Making* by Madinach, Honey, and Light, 2006. Copyright © 2006. Reprinted with permission from the authors.)

In contrast to Bernhardt's framework, this model does not categorize types of data utilized in the process, nor is it specific about what information the intersection of this data may yield to support decision making. The framework demonstrates data-driven decision making as data flows through various processes across various levels of the organization while simultaneously incorporating feedback at all process levels ultimately leading to a decision followed by its implementation. This framework illustrates Senge's concept of systems thinking in that the process is continuous, occurs across levels, and employs feedback loops.

Marsh, et al. (2006) developed a framework adapted from Mandinach, et al. (2006) research (as cited in Marsh, et al.). In similarity to Bernhardt's (2007) framework, Marsh's model recognized that decisions may be informed by multiple types of data shown in Figure 3. The types including outcome, input, process, and satisfaction correspond with Bernhardt's categories. This data yields information following its analysis that must be acted on (actionable knowledge) to inform a type(s) of decision(s).

The types of decisions noted in the framework by Marsh, et al. (2006) are one of six types: set and assessing goals; address individual or group needs; evaluate effectiveness of practices; assess whether client needs are being met; reallocate resources in reaction to outcomes; and, enhance processes to improve outcomes. These decisions can be categorized into two basic types including decisions that identify needs and/or goals, and decisions that use data to act. Following a decision, the process then begins again to determine its effect, thus establishing a continual cycle of evaluation for continuous improvement.

The categorization of data establishes a foundation on which data can be analyzed for interactions within a school at multiple levels for the purpose of making data-driven decisions. For example, using information on instructional processes that students experience may indicate to school principals what programs or instructional strategies are working. One teacher may have students that score consistently well on a particular portion of a standardized test while fellow teachers' students score poorly. This may suggest a need for others to use the same strategy when teaching the concept, or use the same assessments or adjust their lessons, etc. It indicates that this teacher is doing something differently and students are experiencing success

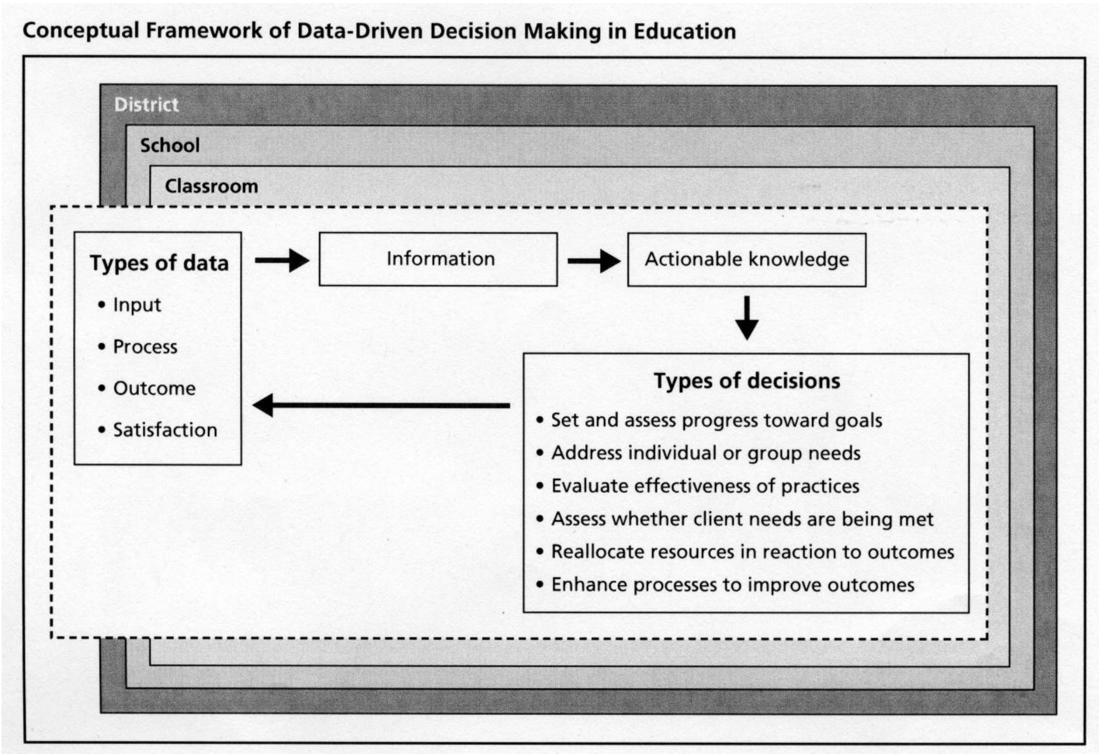

Figure 3. Conceptual framework of data-driven decision making in education. Data from varying sources affects information and actionable knowledge, which affects the type of decision made. From *Making Sense of Data-Drive Decision Making in Education* by Marsh, Pane, and Hamilton, 2006, Santa Monica, California, The RAND Corporation. © The RAND Corporation. Reprinted with permission.)

backed by hard data. Each of these areas of measurement can be used on an individual basis. However, Bernhardt (2000) noted that often clearer pictures of what is occurring in a school or district are provided when these areas are used together.

The framework developed by Marsh, et al. (2006) illustrates that data-driven decision making must be viewed across levels of the educational system because types of decisions and data collected will vary. It shows that conditions at each level may affect the process as well. For example, accuracy and accessibility to data that may exist at the school level may not exist at the classroom level. This absence will influence the ability to turn data into actionable knowledge. Although it is not depicted in the figure, Marsh, et al. (2006) stated that any step may result in the need for additional data before proceeding.

Types of Data Schools Collect

Input, process, outcome, and satisfaction are all types of data identified by Marsh, et al. (2006) that research has indicated schools collect (Mandanich, et al., 2006; U.S. Department of Education, 2008). The need to utilize multiple forms of data for informed decision making is well documented (Breiter & Light, 2006; Halverson, Grigg, Prichett, & Thomas, 2005; Marsh, et al., 2006). Input data consists mainly of demographic data on items such as enrollment, attendance, grade level, ethnicity, gender, home background, and language proficiency. All of these data sub-categories are required reporting components of NCLB in one form or another for students, but many schools also collect a variety of demographic data from other stakeholders. This data also includes demographic data on teachers, principals, and parents (U.S. Department of Education, 2008). School process data refers to programs, instructional strategies, and classroom practices. Satisfaction data can be collected to help determine the beliefs of the schools' stakeholders regarding the school system. Actions of stakeholders are often based on a perception, which heightens the importance of perception data collected through the use of questionnaires, surveys, interviews, and observations. Outcomes data refers to the collection of data related to academic performance. These data items include standardized test results, grade point averages, standard assessments, teacher observations, authentic assessments, graduation rates, formal and informal assessments, and dropout rates (Bernhardt, 2007).

Research has shown that principals use multiple data sources to inform decisions, but not multiple data types. Research by Englert, et al. (2004) showed that over 80% of 124 principals surveyed reported using 5 or more of 18 possible data sources regularly to inform decisions. The five most used types of data included: attendance rates (84%), parent/community feedback (75%), district created assessments (70%), teacher observations (69%), and scores on standardized tests other than state assessments (65%). Three of these five data sources are considered outcomes type data. When principals were asked to identify what data they used to monitor changes in growth and student achievement in their schools, 90% responded "percentage of students proficient by grade level."

Use of Data in Schools

Several studies have found that data is most frequently used in schools by teachers and principals to inform decisions related to improving processes to improve outcomes as measured by NCLB or other outcomes data (Dembosky, et al., 2005; Stecher & Hamilton, 2006; United States Department of Education, 2008). According to O'Day, Clune, Mason, and Thiel (1999), Milwaukee Public School Reform showed that data-driven decision making was earmarked as integral to the process of school improvement as measured through state tests. Teachers and administrators used data to identify problems and make a variety of data-driven decisions in several areas including: for identifying gaps in learning for individual and subgroups of students, schools, and districts; for developing curriculum, instruction, and professional growth; for setting goals for school improvement; for monitoring student academic progress; for informing parents; for tracking individual test scores; for placing students; and, for identifying promising practices.

In Pennsylvania, standardized assessment data is provided to principals in schools identified as failing in a preformatted plan for school improvement that mirrors data-driven frameworks discussed with the exception that standardized score data is expected to be utilized. Over use of standardized score data to make decisions may be producing undesirable results because teachers, at the direction of principals, are spending significant time teaching topics emphasized on the tests, styles and formats used in the tests, and on test taking skills. This concept was defined by one Pennsylvania teacher as, "teaching toward the test" (Stecher & Hamilton, 2006, p. 8). Over-use or single use of one data measure to make decisions is considered bad practice (Bernhardt, 2007; Koretz, 2003; Senge, 2000; Streifer, 2004).

In addition, the practice of relying so heavily on state test score data has spawned the phenomenon of *bubble kids* as noted by Marsh, et al. (2006). Bubble kids are created by singling out and devoting more resources to students on the border of scoring proficient state assessments under the assumption that they will be successful if provided extra help. The scores of these students are typically 5 to 10 points from the cutoff for proficiency. Light, Wexler, and Heinze (2004) found in their study of 146 principals and 1,400 teachers that a majority of principals and teachers reported the use of data in this way to support decisions to reallocate resources in an effort to raise test scores by providing these targeted students access to pull outs, after school, or special programs to raise their score up one level.

Principals have reported increased use of data to inform all types of decisions (Stecher & Hamilton, 2006; U.S. Department of Education, 2008). However, great variability exists among schools as well as within schools in their reported forms and level of sophistication in using data. Marsh, et al. (2006) stated that Pennsylvania schools and districts appeared to be at the very early stages in their development and use of data in comparison to other states.

Stecher and Hamilton (2006) found in a survey of principals in Pennsylvania, Georgia, and California that most principals reported data from standardized state tests as useful in making various decisions encompassed by frameworks noted previously. Their study indicated that this single type of data was used by most principals to inform decisions of several types including: making changes to curriculum and instructional materials; developing a school improvement plan; identifying students in need of additional instructional support; making decisions on how much time is spent on each academic subject; and, focusing teacher professional development. Fourteen percent of principals supported the use of the data in identification of teacher strengths and weaknesses. Further research by Englert, et al., (2004) supported these findings by holding that most principals use this data to evaluate effectiveness of practices, programs, and identify needs for professional development.

Data Management and Analysis Systems

Student achievement data including student grades and scores on state tests are among the most frequently electronically stored data elements in schools (United States Department. of Education, 2008). From 2005 to 2007 there was a 20% increase in schools reporting this electronic capability. The pursuit of continuous improvement in schools has given rise to the use of data management and analysis systems (DMAs). These are prepackaged software systems that deliver updated analyses and comparisons of student data including demographics, standardized test scores, grades, local assessment scores, as well as methods for continued collection such as online testing that is automatically and immediately scored (Stein, 2003). For many schools, data-driven decision making has created a problem in that DMAs require a significant fiscal investment and ideological commitment on behalf of schools. A commitment to DMAs cannot be afforded by some schools and some are not ready to make that commitment, which puts them at a disadvantage (Stein, 2003). In the absence of DMAs, data collection can be inconsistent, time consuming, and inaccurate. Consequently, the use of data and its analysis can be tainted, therefore, leading schools toward ill advised decisions.

Pennsylvania Data Analysis Tools

Whereas most teachers reported having access to electronic student data systems, the proportion of teachers with data system access that also have access to tools for making instructional decisions informed by data remains relatively low (U.S. Department of Education, 2008). The state of Pennsylvania encourages the use of seven such data analysis and management tools for use by the public and educators. These tools support the state's *Getting Results* framework, which requires all schools identified as in need of improvement per NCLB to utilize this data-informed decision framework that mirrors the framework established by Marsh, et al. (2006).

Research suggested that the majority of Pennsylvania school principals are using the data tools provided to them by the Pennsylvania Department of Education (McCaffrey & Hamilton, 2007). Pennsylvania categorized these tools as *system-level* tools (accessible by the general public) and *student-level* tools that are available to educators. System level tools, including Pennsylvania Adequate Yearly Progress (PAAYP), SchoolDataDirect, and the National Assessment of Educational Progress (NAEP), are web-based platforms that enable users to view school performance over time categorized by school, district, and standard. These tools provided the ability to measure school and district progress against the requirements of NCLB. McCaffrey and Hamilton (2007) found that over 60% of principals indicated that they found information in reports like those generated by the PAAYP website that shows school wide results disaggregated by subgroup to be “very useful” in their decision making. SchoolDataDirect and NAEP provide additional data including school district revenue, taxes, and community demographics.

Pennsylvania’s student level data tools are web-based, and include PSSA Data Interactive (Emetric), 4Sight Benchmarks, and the Pennsylvania Value-Added Assessment System (PVAAS). The Emetric system is designed to provide access to student performance results on state standardized tests by summary or individual student, content, statistics, aggregation levels, subgroups, score variables, or in any combination. McCaffrey and Hamilton (2007) found that over 70% of Pennsylvania principals indicated that data reports that included school wide disaggregated standardized score results by topic or skill were “very useful” in their decision making process. Though the Emetric system generates reports of this type, only 31% of principals reported the Emetric system as “very useful,” which suggests that many Pennsylvania principals may lack awareness regarding Emetric reporting capabilities. The Pennsylvania Department of Education has established the Pennsylvania Benchmark Initiative using 4Sight Benchmark exams that are aligned with state standardized tests to provide an estimate of student performance on the PSSA, guide instruction, and identify gaps in student learning. These results can be uploaded and analyzed at the Success for All website.

The use of these types of tools is a characteristic of schools that are effective in data-driven decision making as measured by state standardized tests (Armstrong & Anthes, 2001;

Breiter & Light, 2006; Wayman, Stringfield, & Yakimowski, 2004). Principals reported a high frequency in use of these types of tools to inform parents, identify student needs, and in discussion with teachers regarding test scores (Mandinach, et al., 2006). The Grow Network web-based tools have been shown to be effective in the practice of informing instruction (Breiter & Light, 2006; Mandinach, Rivas, Light, Heinze & Honey, 2006). Light et al. (2004) found that data tools like Grow Network tool have helped principals and teachers turn data into information that can be used to inform decisions and improve student learning as measured by standardized assessments.

Value-Added Modeling

The term value-added is an economic concept that defines an organization's effectiveness by increases in its productivity measured over time (Meyer, 1997). Applied to education, the term *value-added* or *value-added modeling* (VAM) is a collection of complex statistical analyses employing models that are designed to interpret standardized test scores and other factors related to student achievement to determine and to project academic growth (McCaffrey, 2003; Rubin, Stuart, & Aanutto, 2004). Data analysis conducted using value-added modeling provides accountability for schools that recognizes the individuality of student learning by holding schools accountable for measureable gains for each student over a school year as opposed to a predetermined proficiency score.

Pennsylvania Data Analysis Tool for Value-Added Assessment

The Pennsylvania Valued-Added Assessment System (PVAAS) is a data analysis tool unlike others available to principals. At the start of 2006, all Pennsylvania schools received reporting through PVAAS based on William Sanders'(1994) value added model (Pennsylvania

Department of Education, 2008). PVAAS is a statistical analysis of PSSA assessment data so that it may be compared longitudinally to measure schools' influence on the academic progress rates of groups of students from year to year. This web-based tool is designed to yield two types of data for schools: data on the gain or growth of cohorts; subgroups of students; or both as well as individual student projection data that established a prediction of future performance on the PSSA (Pennsylvania Department of Education, 2008). The PVAAS system provides feedback to the following questions to help educators make data informed decisions:

1. Did each cohort (grade level) make a year's worth of growth in reading (grades 4-8 and 11), mathematics (grades 4-8 and 11), science (grades 4, 8 and 11), and writing (grades 5, 8, 11)?
2. Did each subgroup of students make a year's worth of growth in reading (grades 4-8 and 11), mathematics (grades 4-8 and 11), science (grades 4, 8 and 11), and writing (grades 5, 8 and 11)?
3. Is each individual student on a trajectory to reach proficient or advanced levels on a future PSSA reading, mathematics, science, and writing assessment? (Pennsylvania Department. of Education, 2008, p. 3)

The PVAAS platform was based on the fundamental assertion in value-added modeling that schools are accountable for students making a year's worth of academic growth in a year's worth of schooling (Carey, 2004; Hershberg, Simon, & Lea-Kruger, 2004). Though PVAAS provides a different analysis of standardized score data, McCaffrey and Hamilton (2007) found in a survey of 411 Pennsylvania principals in PVAAS pilot districts that 28% of principals did not use the tool or were not aware of its availability. Further findings indicated that the PVAAS tool filled a perceived lack of data on student growth for principals that used it because it provided additional information on student achievement. However, McCaffrey and Hamilton pointed out that:

> As long as the primary accountability measures involve only proficiency status, teachers' focus on proficiency and information about proficiency may override usage of value-added data. (p. 86)

In 2009, the PVAAS data tool system was made available to all school districts in Pennsylvania.

Systems Theory

Hammond (1997) recognized Bertalanffy as the originator of general systems theory. According to Hammond, Bertalanffy explained systems as composed of elements that were in continuous interaction. Hammond explained that:

> Bertalanffy characterized organized systems at all different levels were those of wholeness, organization, and dynamic multivariate interaction, where the behavior of the parts was different when studied in isolation than when in the context of the whole. (p. 141)

Educational organizations at the district, building, and state level can all be considered systems on their own and sub-systems of each other. A phenomenon of every system is interdependence or the concept that change in one part of the system will result in change in another part or parts of the system (Littlejohn, 1983). The ability of principals to use data to better understand the interaction of variables (e.g., curricula, teaching practices, student performance, and teacher effectiveness) within a system (e.g., school building, school district, or entire grade level) is linked to evaluation of their school by standardized test scores. For example, Lane, et al. (2003) noted that variations in teacher attitudes and beliefs regarding state standardized assessments, teacher classroom practices, and assessment practices were found to explain variability in schools' performance measured over time on standardized assessments. Each of these variables noted by Stone and Lane (2001) are system attributes from the schools where their data was gathered. Their research examined the effects of variations in multiple attributes and their relation to each other on their effect on the system's or school's performance measured by standardized test data. Bernhardt (2007) as well as Mandinach, et al. (2006) stressed the importance of recognizing the concept that decisions made to affect system performance may be informed by a multiple of data types and require analysis of their interaction within the scope of the whole.

Senge's Systems Theory

Senge (1990) stated that systems thinking theory is the cornerstone component to organizations that learn collectively. According to Senge (1990), learning organizations are:

> Organizations where people continually expand their capacity to create the results they truly desire, where new and expansive patterns of thinking are nurtured, where collective aspiration is set free, and where people are continually learning to see the whole together. (p. 3)

Senge (1990) contended that the organization that learns is one whose members effectively use five fundamental disciplines: personal mastery; shared vision; mental models; team learning; and, systems thinking. To demonstrate *personal mastery*, individual members of an organization must be committed to continual learning, development, and deepening of personal vision in relation to the organization's vision, and understanding their own needs for personal growth. *Mental models* are the unspoken rules of the organization as to how things are done. *Team learning* is the ability of organization members to cooperate as a team to produce desirable results. *Systems thinking* refers to the ability to look at the organization as a whole and understanding the interactions of its parts.

Senge (1990) stated that often simplistic frameworks for problem solving are applied to complex systems. Consequently, focus in solving problems tends to be on the parts of a system or organization rather than the whole. Therefore, the multivariate interactions of the parts as described by Bertalanffy go unrecognized in the context of the system as a whole. He noted that neglecting a systems thinking approach often results in decisions that produce short-term results, but may have devastating long-term consequences. For example, to save money in the short term a company might elect to cut funding for research development. However, in the long-term, the lack of product development through research could ruin the company. Similarly, in the context of education, program cuts that save immediate dollars may have long-term consequences in student performance.

Senge (1990) noted, “The systems viewpoint is generally oriented toward the long-term view” (p. 92). He argued that orientation toward a long-term view is required due to the nature of feedback on decisions within most organizations. Senge (1990) noted that, though people learn and act based on experience, decisions rarely directly affect decision makers. Without a systems approach, educators fail to understand the ramifications of these decisions because there is no experiential feedback. Senge explained that feedback on decisions in an organization is often delayed because the interaction of many variables delays the result. Feedback may, in the short-term, appear positive, but may be negative in the long-term which illustrates Senge’s (1990) contention that feedback loops both for the short and long-term are important. It is a fundamental axiom to systems thinking. Senge argued that influences or variables in systems are in a state of interaction such that one variable does not cause or influence the other. He stated the variables interact with each other as “every influence is both cause and effect”(p. 75). This duality is the distinction between linear thinking and systems thinking. The linear approach holds that variable A causes variable B in a straight line, but the systems approach views these in a circular pattern of causality.

Senge’s Levels of Systems Thinking

Senge (2000) elaborated his theory of systems thinking by noting that it requires the user to see four simultaneously operating levels within a system when problem solving. These levels include events, patterns and trends, systemic structures, and mental models. Each of these levels has one or more questions of focus. Senge likened this to viewing an iceberg in that only what is above the surface of the water can be seen (events), but there is much more at play beneath the surface of the water (patterns, systemic structures, mental models). Events are directly observable actions and behaviors. Patterns emerge as actions and behaviors repeated over time. Systemic structures show the relationship between the patterns. Mental models are the deep seated beliefs and values that define the systemic structures in place. Senge described that at the event level system thinkers attempt to understand what happened. Next, they seek to understand underlying patterns or trends by researching if this event has happened before. At the systemic structures level, system thinkers attempt to determine what variables influence these patterns. Last, at the mental model level, Senge (2000) asserted that system thinkers must practice meta-cognition by asking themselves, “What about our thinking allows this situation to persist?” (p. 80).

Hawley (1998) noted the applicability of the systems approach to address the dynamics and challenges of school improvement. To determine a starting point(s) for improving performance of a system (change), data gathering and analysis are required (Littlejohn, 1983).

Data gathering and analysis provide information that explain how attributes interact and comprise a system. Knowledge of these interactions provides a foundation on which decisions effect desired changes in the system (Littlejohn, 1983). Data gathering and analysis are, therefore, a critical component to making decisions in schools. This process has been termed *data-driven decision making*. Streifer (2004) termed the process "data guided decision making" (p. 9) by noting in similarity to Senge's (1990) systems thinking that questions involving data, especially related to education, are often complex involving the interaction of many variables. As a result, decisions are not often purely driven by data; the data is used to inform decisions.

Summary

The review of literature of data-driven decision making and NCLB evidenced several points pertaining to educational practices regarding the use of data in schools. The collection and use of data in schools has increased sharply since the passage of NCLB. This influence is easily indentified at the school level as schools strive to meet the accountability mandates of NCLB. This has made assessment data in both standardized and local formats the most used type of data in schools. Schools collect vast amounts of data in addition to assessment data that is used by principals to inform different types of decisions in many contexts. This process has given rise to new tools like value added modeling and web based software designed to analyze standardized test scores for school improvement. Effective data-driven decision making occurs in a culture where expectations for data use are clearly established for the purpose of continuous improvement, which implies regular data informed decision making of varying types and regular use of data tools.

CHAPTER 3

PROCEDURES

Introduction

School districts are held accountable for progress as measured on standardized assessments required by NCLB. Increased demands for school accountability under NCLB have spawned a thrust for the collection and analysis of data in schools as part of a new avenue for school reform termed “data-driven decision making” (Education Commission of the States, 2002). Student test scores on standardized assessments like the PSSA are one element of student outcome data that receives the most systemic attention by educators to inform process improvement decisions (Englert, et al., 2004; Marsh, et al., 2006). A heightened use and reliance on this state mandated test data is found in schools not making AYP, whereas schools achieving AYP status indicate the use of many types of data in their decision making processes (Marsh, et al., 2006). The process of data-driven decision making in schools has become increasingly important in school improvement efforts. Research has indicated that school principals play a critical role in the data driven decision making process that impacts student achievement (Leithwood, et al., 2004; Marzano, 2005).

Mathematics education has been a contentious topic in the United States as American students have continued to struggle in mathematics proficiency when compared to students in other industrialized nations. Student scores on the National Education Assessment of Progress reflect virtually zero improvement in mathematics scores at the secondary level over the past 37 years. While Pennsylvania students have scored slightly above average on the NEAP assessment, the percentage of Pennsylvania students scoring proficient or above in mathematics on the PSSA in grade 11 is significantly less when compared to the average scores of students in grades 3 through 8. As requirements for proficiency thresholds continue to increase for mathematics under NCLB, school principals are feeling the pressure to increase student achievement in mathematics. This study was designed to explore factors related to secondary principals' use of data and the impact it may be having on the percentage of grade 11 students achieving proficiency in mathematics as measured by the PSSA assessment in the 2008-2009 school year.

Purpose and Research Questions

The purpose of this chapter is to provide a description and rationale for the research design selected for this study. Specifically, this chapter describes the purpose of the study, the research questions, research design, participant selection, instrumentation, data collection, and data analysis procedures. The results of this study are based on the analysis of a survey designed to specifically address the four research questions posed in Chapter 1:

1. What types of data do secondary principals in high and low performing schools use and perceive as effective for increasing student mathematics performance on standardized assessments?
2. What types of data tools do secondary principals in high and low performing schools use and perceive as effective to enhance mathematics scores on standardized assessments?
3. To what extent do secondary principals in high and low performing schools use data systems to enhance student achievement in mathematics on standardized assessments?
4. How does secondary principals' use of data driven decision making in high and low performing schools differ?

a. Is there a difference between secondary principals' use of data types in high and low performing schools?
b. Is there a difference between secondary principals' perception of data type effectiveness for enhancing mathematics scores in high and low performing schools?
c. Is there a difference between secondary principals' use of data tools in high and low performing schools?
d. Is there a difference between secondary principals' perception of data tool effectiveness for enhancing mathematics scores in high and low performing schools?
e. Is there a difference in the extent that secondary principals in high and low performing schools use data systems to enhance student achievement on standardized mathematics assessments?

Population and Sampling

The sample population for the formal study was limited to secondary principals in public school districts in the state of Pennsylvania. Secondary principals in Pittsburgh Public School District and the School District of Philadelphia were excluded from the sample due to their respective school districts' required procedures for participation. A purposeful sample was drawn from the remaining population of 548 secondary principals. The PSSA assessment has four scoring categories that include advanced, proficient, basic, and below basic. The Pennsylvania Department of Education calculates overall proficiency ratings for each school by tallying the percentage of students scoring in the advanced and proficient categories. Grade 11 student assessment data for the PSSA mathematics test given in the 2008-2009 school year was analyzed to numerically rank the 548 public secondary schools based on the total percentage of grade 11 students that scored proficient and advanced in each school. Principals in schools ranked 1 through 150 and 398 through 548 were purposefully selected and invited to participate in the study. Through the analysis of survey responses, this research determined whether principal engagement and perceptions related to the use of data systems varied across schools. An understanding of the relationships among high and low performing schools will support the development of strategies for school improvement. This research on the use of data systems may assist principals and teachers in developing strategies for improving student achievement in mathematics by using data.

School principals invited to participate in this study were identified from a current list of schools archived on the Pennsylvania Department of Education's web site. The list contained published PSSA assessment data in mathematics for all schools at all levels disaggregated by student demographics as required under NCLB. Principal contact information was obtained from schools or acquired from school websites. All necessary confidentiality agreements were in place prior to the distribution of or use of any information obtained from schools or school websites.

A letter was mailed to the principal of each school meeting the criteria to participate in the study. The correspondence included a description of the proposed study and directions for completing the study online. In addition, a two dollar bill and note card were enclosed with each letter. The note card contained a brief description about the history of the two dollar bill being used to celebrate American public education on one side. The two dollar bill was enclosed to encourage participation in the study.

Instrumentation

This research utilized a survey designed to gather perceptions of practicing school principals regarding the use of data to improve mathematics scores on standardized assessments in their schools. The survey was created in two parts. The first part of the survey collected relevant demographic information from principals including: gender; age range; years worked in current school; and, school district size. The second part of the survey consisted of 41 items to determine what relationships exist among principals' use of data systems, use and perceived effectiveness of data tools and types in enhancing student achievement on standardized mathematics assessments. The items for the second part of the survey were created based on a review of previous instruments and frameworks for data driven decision making found in the literature (Bernhardt, 2007; Englert, et al., 2004; Goodwin, et al, 2003; Marsh, et al., 2006; Means, et al., 2007).

A list of specific data types used to increase student achievement was generated based on the frameworks for data driven decision making developed by Bernhardt (2007) and Marsh, et al. (2006) and the instrument used by Means, et al. (2007). The compiled list contained 21 different kinds of data that could be categorized into four specific types including input data, outcomes data, process data, and satisfaction data. Principals indicated yes or no if they used each kind of data in the list to enhance student mathematics achievement on standardized assessments. A follow up question was asked for each specific type of data that principals indicated they used. The follow up question asked principals to rate the effectiveness of each data type in enhancing student achievement in mathematics as measured by standardized assessments on a Likert scale of 1 (Not Effective) to 5 (Very Effective).

A total of 13 items designed to determine the extent which principals engage in the use of data systems to improve mathematics scores were adapted from a previous instrument developed by Englert, et al. (2004), and were included in the original instrument for this study. In a study sponsored by the Mid-Continent Research for Education and Learning (MCREL), Goodwin, et al. (2003) reviewed the literature on accountability systems for the purpose of designing an instrument to measure effective accountability systems. Their review identified 12 frequently cited characteristics of good systems that are essential to examine when evaluating accountability systems. Seven of these characteristics were determined to be most relevant by an expert panel of principals including: high expectations for all students; high-quality assessments aligned with standards; alignment of resources, support, and assistance for improvement; sanctions and rewards linked to results; multiple measures; diagnostic uses for data; and, readily understandable to the public. After the survey was pilot tested for content validity with a group of former principals and school administrators, wording in some items was modified and additional items were added. The final survey was first used by Englert, et al. (2004), and sent to 308 principals in four states including Colorado, Kansas, Missouri, and South Dakota. State representatives from each of these states were asked to identify 20 school districts that would represent the state in terms of percentage of low, middle, and high performing districts; percentage of rural, suburban, and urban districts; varying amounts of per-pupil expenditures; and varying percentages of minority students. A total of 121 principals in 48 districts completed the survey for a response rate of 39%.

Of the seven characteristics identified as essential to accountability systems, principals' "diagnostic uses for data" is defined as "using data to diagnose problems and work toward solutions" (p. 4). The use of data in this way was identified as a potentially successful use of a data system to increase student achievement (Bernhardt, 2007; Streiffer, 2004). The survey used by Englert, et al. (2004) had 13 items to measure the extent of principals' "diagnostic uses of data" in high and low performing schools as measured by student performance on state standardized assessments. Participant responses were recorded on a Likert scale ranging from 1 (To no extent) to 5 (To a great extent). Analysis revealed that the 13 items on the Englert, et al. instrument used to measure principals' "diagnostic uses for data" were determined to have a .88 Cronbach's Alpha rating for reliability. Englert, et al. found that principal responses in this category showed a significant difference when compared in high and low achieving schools as measured by state standardized assessments. For the purposes of this research, these items were, therefore, determined to be suitable for adaption to measure principals' use of data systems to increase student achievement in mathematics.

The 13 items from the Englert, et al. (2004) instrument were used to measure principals' use of data systems to enhance student mathematics achievement. Two additional items were added by the researcher based on the review of literature. Each of the two added items sought to evaluate the extent of principals' engagement in practices for using data systems including the use of multiple data sources to evaluate student progress, and using data to differentiate instruction (Bernhardt, 2007; Schmoker, 2006). The wording and scaling of the items adapted from the Englert, et al. (2004) instrument were modified for their use in this research with written permission from MCREL. Wording was changed to reflect that principals were being asked to rate the extent that the element in the item was used or done for the purpose of enhancing mathematics scores on standardized assessments. Participants in this study were asked on these items to rate the extent that they engaged in various uses of data systems to enhance student achievement in mathematics for these 15 items on a Likert scale of 1 (To No Extent) to 5 (To a Great Extent).

A list of specific data tools was developed through the review of literature relating to the use of data tools in schools. A list of data tools supported for use by the Pennsylvania Department of Education was obtained from the PDE website. The list also included a description generated by PDE for each data tool. This list was supplemented with four additional tools identified from the literature (Wayman, Stringfield, & Yakimowski, 2004). Principals were asked to select yes or no to indicate if they used each data tool in the list to enhance student mathematics achievement on standardized assessments. A follow up question was asked for each data tool that principals indicated they used. The follow up question asked principals to rate the effectiveness for each data tool they used in enhancing student achievement in mathematics as measured by standardized assessments on a Likert scale of 1 (Not Effective) to 5 (Very Effective).

Validity

The researcher established content validity of the survey using input from a panel of experts that consisted of six school principals in Pennsylvania who have experience using data in efforts to enhance student achievement on the PSSA mathematics assessment. The work experience of the panel members in public education exceeds 140 years. Following is a list of the expert panel members (school principals) by school district:

Somerset Area School District: Dr. Mark Gross;

Clarion Area School District: Dr. Randy Cathcart;

Blacklick Valley School District: Dr. Luke Lansberry;

Spring Cove School District: Dr. Dave Crumrine; and, St. Marys Area School District: Mr. Joseph Schlimm and Mr. James Wortman.

Input was gathered from the expert panel using an item analysis of the instrument. Each of the panel members received an email inviting them to complete the item analysis. The panel members were asked to indicate whether each "data type" item and each "data tool" item contained either "vital information," "adequate information," or "not much here." Panel members were also asked to indicate whether each question in the survey section related to principals' use of data systems contained either "vital information," "adequate information," or "not much here." Any questions that were identified by three or more expert panel members as "not much here" were eliminated. Four items (#12, #18, #20, #25) from the section of the survey on data types were recommended by panel experts as "not much here." Two questions (#31 and #34) from the section of the survey on principals' use of data systems were also recommended by panel experts as "not much here." One question (#52) from the section on data tools was eliminated. Questions 2-6 in the pilot study pertained to demographics. The results of the pilot study for content validity are listed in Table 1.

The feedback from the panel of experts indicated that the instrument designed for this study has content validity for measuring principals' use of data tools, data types, and data systems to enhance student achievement in mathematics as measured by standardized assessments. The minimum percentage of votes in the Adequate to Vital Information range was 67%. Seven items (#12, #18, #20, #25, #31, #34, #52) received a percentage of votes ranging from 50% to 100% in the Not Much Here range. These items were eliminated. Two items received 100% of votes in the Vital Information range. A total of 81% of the survey items

Table 1

Results from Content Validity Pilot Study of Instrument

Question Number	Percentage of Expert Votes Not Much Here	Percentage of Expert Votes Adequate	Percentage of Expert Votes Vital Information
7	17%	50%	33%
8	0%	83%	17%
9	0%	83%	17%
10	0%	83%	17%
11	0%	33%	67%
12	83%	17%	0%
13	0%	50%	50%
14	0%	50%	50%
15	0%	17%	83%
16	0%	83%	17%
17	17%	83%	0%
18	67%	33%	0%
19	0%	50%	50%
20	83%	17%	0%
21	0%	83%	17%
22	0%	17%	83%
23	0%	67%	33%
24	0%	83%	17%
25	83%	17%	0%
26	0%	83%	17%
27	0%	83%	17%
28	0%	33%	67%
29	0%	17%	83%
30	0%	33%	67%
31	67%	33%	0%
32	0%	17%	83%
33	0%	17%	83%
34	50%	50%	0%
35	0%	33%	67%
36	0%	50%	50%
37	0%	17%	83%
38	0%	17%	83%
39	0%	17%	83%

Table 1 Continued

Results from Content Validity Pilot Study of Instrument

Question Number	Percentage of Expert Votes Not Much Here	Percentage of Expert Votes Adequate	Percentage of Expert Votes Vital Information
40	0%	0%	100%
41	0%	17%	83%
42	0%	17%	83%
43	0%	17%	83%
44	0%	50%	50%
45	0%	50%	50%
46	0%	46%	33%
47	0%	83%	17%
48	0%	33%	67%
49	0%	17%	83%
50	0%	17%	83%
51	0%	0%	100%
52	100%	0%	0%
53	0%	50%	50%

received 100% of the votes in the adequate to vital information range which demonstrated that the instrument has content validity for evaluating principals' use of data types, data tools, and data systems to enhance student achievement in mathematics.

Reliability

Formal piloting of the survey instrument was conducted in April of 2010. The pilot population consisted of a convenience sample of 15 principals working in Pennsylvania. The formal pilot study was conducted to establish reliability for the survey questions. Practicing secondary school principals were chosen as participants for this study because of their job experience which includes the collection and use of data to assess their respective schools' performance in mathematics under NCLB. Secondary principals in seven schools that had mathematics scores ranking between 150 and 200 were chosen to form the upper performance group for the pilot study, while eight secondary principals in schools where mathematics scores ranked between 348 and 398 formed the lower performance group. A response rate of 93.3% was obtained for the pilot study with seven principals in both the upper and lower group responding for a total 14 of 15 principals completing the entire survey in the pilot study.

Participants were asked to select data types and data tools that they used to improve mathematics scores on PSSA exams. They were then asked to rate the extent that the type of data or data tool chosen was effective in enhancing mathematics scores in their school by rating the items on a Likert scale from 1(Not Effective) to 5 (Very Effective). Principals' use of data systems was also evaluated using 13 items that required participants rate the extent of their use of each item on a Likert scale from 1 (To No Extent) to 5 (To a Great Extent). Correlations were used to assess the reliability of the instrument to establish Cronbach alpha levels greater than .80. A Cronbach's alpha level of .88 was established for the survey sub-category, use of data systems.

Other sub-categories numbered too few cases to establish a Cronbach's alpha level in the piloting phase of the study.

Data Analysis

Analysis of variance and chi-square were used to analyze the data. Questions on the instrument were grouped into eleven sub-categories including: use of input data; use of process data; use of outcomes data; use of satisfaction data; perceived effectiveness of input data; perceived effectiveness of process data; perceived effectiveness of outcomes data; perceived effectiveness of satisfaction data; use of data systems; use of data tools; and, perceived effectiveness of data tools used. The sub-categories represent a set of questions grouped together based around a common content focus. A one-way ANOVA was used to compare principal responses by sub-category for the upper and lower ranked schools chosen for the study.

A two-way ANOVA was conducted to determine whether significant differences existed between the mean scores for principals in the upper schools, lower schools, and principal age. Then a two-way ANOVA was completed to determine if significant differences existed between the upper schools, lower schools and years of working as the principal at the school. Additionally, a two-way ANOVA was conducted by the researcher to analyze if significant differences existed between the mean scores of principals in the upper schools, lower schools, among rural, suburban, and urban schools.

A chi-square was completed to determine any significant differences between the principal responses in the upper and lower schools regarding their use of input data, outcomes data, process data, satisfaction data, and use of data tools. A second chi-square analysis was completed to determine whether significant differences existed between the upper and lower schools, gender, and these survey items. An additional two-way ANOVA was conducted to examine differences relating to gender and the eleven subcategories.

Summary

Secondary principals chosen to participate in this study ranged from those working in schools with highest proficiency percentage to the lowest proficiency percentage for grade 11 students on the 2008-2009 PSSA mathematics assessment. Secondary principals were selected based on numerous findings that secondary students lag behind elementary and middle school students in mathematics achievement, since the passage of NCLB. Additionally, American secondary school students have not performed as well on standardized assessments in mathematics as their counterparts in other industrialized nations. The pressure to improve has made the use of data systems to increase standardized test scores a necessary skill for principals.

It was the purpose of this study to analyze principals' use of data systems including various data types and tools to gain an understanding of what principals' in higher performing schools do differently with data as compared to principals in lower performing schools. An instrument of original design constructed by the researcher with questions adapted from Englert, et al. (2004) was used to gather data. The data was analyzed using one-way and two-way ANOVA and chi-square analysis in order to answer the research questions about the relationships between principals' perceptions and use of data tools, data types, and data systems to enhance mathematics scores on standardized assessments in high and low performing schools.

CHAPTER 4

DATA AND ANALYSIS

Introduction

This study was designed to compare secondary principals' use of data systems, perceptions of data tools and data type effectiveness in high and low performing schools as measured by the percentage of grade 11 students rated proficient on the 2008-2009 PSSA standardized mathematics assessments. The instrument was created by the researcher and included 11 items from Englert's (2004) study on principals' use of data that were adapted to survey principals' use of data systems to enhance mathematics performance in their school. For the remainder of this data review and analysis in this chapter, principal responses are identified as upper for those with performance ranks from 1 to 150 and lower for those with performance ranks from 151 to 300.

The following research questions were addressed:

1. What types of data do secondary principals in high and low performing schools use and perceive as effective for increasing student mathematics performance on standardized assessments?
2. What types of data tools do secondary principals in high and low performing schools use and perceive as effective to enhance mathematics scores on standardized assessments?
3. To what extent do secondary principals in high and low performing schools use data systems to enhance student achievement in mathematics on standardized assessments?

4. How does secondary principals' use of data driven decision making in high and low performing schools differ?
    a. Is there a difference between secondary principals' use of data types in high and low performing schools?
    b. Is there a difference between secondary principals' perception of data type effectiveness for enhancing mathematics scores in high and low performing schools?
    c. Is there a difference between secondary principals' use of data tools in high and low performing schools?
    d. Is there a difference between secondary principals' perception of data tool effectiveness for enhancing mathematics scores in high and low performing schools?
    e. Is there a difference in the extent that secondary principals in high and low performing schools use data systems to enhance student achievement on standardized mathematics assessments?

Analysis focused on 11 sub-categories. Four of the sub-categories related to principals' use of various types of data including use of input data, use of process data, use of outcomes data, and use of satisfaction data. Four sub-categories related to principals' perceived effectiveness of various types of data in enhancing student achievement in mathematics including perceived effectiveness of input data, perceived effectiveness of process data, perceived effectiveness of outcomes data, and perceived effectiveness of satisfaction data. One sub-category related to principals' use of data systems to enhance mathematics achievement. Two additional categories related to principals' use of data tools and perceived effectiveness of data tools in enhancing mathematics achievement.

Sample Population

The sample population for this study included 300 secondary principals in public schools across Pennsylvania. One hundred forty-seven principals completed the entire survey for a response rate of 49%. The researcher first conducted a descriptive analysis of the participant background characteristics. This included frequencies and percentages for gender, age, number of years working as a principal at their current school, school performance group, and whether the school was identified as rural, suburban, or urban. The descriptive statistics for the background characteristics of the entire sample are presented in Table 2. The majority of participants were male (79.6%). In terms of age, 53.7% were 40-54 years old. Years as principal varied with the largest group of participants indicating they had been principal in their school for one-four years (49.7%). District size also varied; more than half of the participants (56.5%) were in districts with 0-1,999 students.

Statistical Analysis

Analysis of variance (ANOVA) and chi-square were used to analyze the data as guided by the research questions. Huck (2004) stated that one-way ANOVA f-tests are an appropriate statistic that may be used to compare means from two groups in order to test a null hypothesis. Gelman (2004) noted the importance and usefulness of the information provided by the ANOVA test. The design of the study is focused on comparison of means for two independent samples. Therefore, the use of one-way ANOVA is appropriate for comparison of means between the upper and lower group of principals.

Table 2

Participant' Background Characteristics

Characteristic	N	%
Gender		
Male	117	79.6
Female	30	20.4
Total	147	100
Age		
25-39 years	41	27.9
40-54 years	79	53.7
55+ years	27	18.4
Total	147	100
Years as a principal in your current school		
1-4 years	73	49.7
5-8 years	42	28.6
9+ years	32	21.8
Total	147	100
School district's size		
Rural (0-1,999 students)	83	56.5
Suburban (2000-6,999 students)	51	34.7
Urban (7,000+ students)	13	8.8
Total	147	100
Performance		
Upper	75	51.0
Lower	72	49.0
Total	147	100

Data Analysis

Chi-square statistic was used to investigate whether distributions of categorical variables differed from one another (Huck, 2004). Huck stated that an independent-sample chi-square test is appropriate when two independent samples are compared with respect to a dichotomous dependent variable. The design of this study collected data that can be analyzed this way by asking principals in two groups to respond yes or no as to whether or not they used an individual data item or individual data tool. Chi-square was selected as a method of analysis for this data.

Two-way analysis of variance is designed to explore the effect of two factors simultaneously (Huck, 2004). Gelman (2004) suggested that the two-way ANOVA holds advantages that include the ability to use a smaller sample size, reduced random variability with the use of more than one factor, and the ability to examine the effects of two factors simultaneously. The data collected in this study are suitable for analysis using two factor ANOVA. Though not specifically guided by the research questions, the two factor effects of the independent variables including performance and age, performance and size of school, followed by performance and years as principal at current school were analyzed using two-way ANOVAs.

Questions on the instrument were grouped into 11 sub-categories including: use of input data; use of process data; use of outcomes data; use of satisfaction data; perceived effectiveness of input data; perceived effectiveness of process data; perceived effectiveness of outcomes data; perceived effectiveness of satisfaction data; use of data systems; use of data tools; and, perceived effectiveness of data tools used. The sub-categories represent a set of questions grouped together based around a common content focus. Analysis of variance was conducted in 11 sub-categories. Chi-square analysis was used to analyze the sub-categories use of data type and data tools. Descriptive statistics were also completed for sub-categories and individual items related to the use of data types, use of data tools, perceived effectiveness of data types, and perceived effectiveness of data tools. Principal responses were compared by sub-category for the upper and lower ranked schools chosen for the study to examine differences. Forty-nine percent were classified into lower performing schools; 51% were classified into upper performing schools (see Table 2).

Use of Data Types

The first research question addressed by the study sought to determine the types of data secondary principals in high and low performing schools use and perceive as effective for increasing student mathematics performance on standardized assessments. Principals responded either yes or no to items that asked if they used a specific type of data in their efforts to increase mathematics scores on standardized assessments. Yes responses were recorded numerically as a 2 and no responses were recorded numerically as a 1 in order to generate means for comparison of principals' use of data types and individual data items within each sub-category. Descriptive statistics on principals' use of data types are presented in Table 3. Data types are ordered from the highest mean to the lowest mean of the upper schools. The results in Table 3 show that principals in both performance groups use different types of data to varying degrees. Outcomes type data appears to be used the most by principals in both groups followed closely by process data. Nearly half of the principals in both groups reported the use of satisfaction data to increase mathematics scores while a relatively low number reported the use of input data. Additionally, Table 3 shows that there is more reported use by principals in the lower group for three of the four data types. A higher number of principals in the upper group reported use of process data in efforts to achieve higher mathematics scores on standardized assessments.

Table 3

Principals' Use of Data Types

Data Type	School	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Min	Max
	Lower	75	1.66	.24463	.02825	1.6085	1.7210	1.00	2.00
Outcomes	Upper	72	1.63	.26770	.03155	1.5661	1.6919	1.00	2.00
	Total	147	1.65	.25594	.02111	1.6055	1.6889	1.00	2.00
	Lower	75	1.56	.32868	.03795	1.4888	1.6401	1.00	2.00
Process	Upper	72	1.61	.33567	.03956	1.5322	1.6900	1.00	2.00
	Total	147	1.59	.33181	.02737	1.5332	1.6414	1.00	2.00
	Lower	75	1.49	.35231	.04068	1.4078	1.5699	1.00	2.00
Satisfaction	Upper	72	1.46	.34685	.04088	1.3768	1.5398	1.00	2.00
	Total	147	1.47	.34879	.02877	1.4171	1.5308	1.00	2.00
	Lower	75	1.30	.33136	.03826	1.2238	1.3762	1.00	2.00
Input	Upper	72	1.26	.32741	.03859	1.1870	1.3408	1.00	2.00
	Total	147	1.28	.32880	.02712	1.2287	1.3359	1.00	2.00

Individual data items that comprise the four data type sub-categories are presented in Table 4 ranked from highest percentage to lowest percentage of principals reporting use. In similarity to the findings of Marsh, et al. (2006), "standardized test score data" and "data collected from local benchmark assessments" are used by a higher percentage of principals in both groups in efforts to increase mathematics scores on standardized assessments. However, fewer principals in the upper group reported using these individual data types. Nearly every principal in the lower group reported that they use "data collected from local benchmark assessments" in their efforts to improve scores. Both groups reported "feedback data from teachers" from the satisfaction data type sub-category as the third highest in principal use. A higher percentage of principals in the upper group reported more use of "data about curriculum needs" than principals in the lower group. It is interesting to note that principals in the lower group reported a higher percentage of use on nearly every individual item in the outcomes sub-category. It appears that both groups may be over using one type of data in their decision making processes.

Data Type Effectiveness

Principals were asked to rate the effectiveness of data types on a five item scale from "not effective" to "very effective." Each item on the scale was assigned a number in the analysis for the purpose of generating a mean effectiveness rating. A rating of "not effective" was given the numerical value of 1. The next rating, interpreted as "slightly effective," was given the numerical value of 2. The next item, interpreted as "somewhat effective" on the scale was assigned the numerical value of 3. The next item, interpreted as "effective," was given the numerical value of 4. The final rating on the scale, "very effective," was assigned the numerical value of 5.

Table 4

Percentage of Principals Indicating Use of Data Type by Individual Item

Item	Data Type	Upper	Lower
Standardized test score data	outcomes	79	81
Data collected from local benchmark assessments	outcomes	76	99
Data about curriculum needs	process	71	60
Feedback data from teachers	satisfaction	71	75
Report card grade data	outcomes	69	59
Teacher observation data	outcomes	67	67
Data about school program effectiveness	process	60	57
Teacher generated authentic assessment data	outcomes	53	54
Data about best practices for instruction	process	53	52
Student attendance data	outcomes	49	57
Value added assessment data	outcomes	46	49
Student socioeconomic data	input	42	53
Feedback data from students	satisfaction	36	44
Data on perceptions of the learning environment	satisfaction	31	28
Student ethnicity data	input	29	24
Student language proficiency data	input	18	21
Student gender data	input	17	21

Note. Upper N = 75; Lower N = 72.

The descriptive results of these mean ratings are presented in Table 5 by data type. Sample sizes vary as principals had to report use of every individual data item within a sub-category in order to calculate an effectiveness rating. Data type sub-categories are listed from highest mean rating to lowest mean rating of the upper performance group. Though the highest percentage of principals reported use of outcomes type data in both groups, principals in the upper group rated process type data as more effective in efforts to increase mathematics scores with a mean effectiveness rating of 4.0. Principals in the lower group reported outcomes type data as most effective in their efforts to increase mathematics scores. Both groups rated input data as least effective. Mean effectiveness ratings for the lower group were similar across the remaining three sub-categories of process, outcomes, and satisfaction type data.

Table 5

Principals' Rating of Data Type Effectiveness

Data Type	School	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Min	Max
	Lower	20	3.93	.56816	.12704	3.6674	4.1992	3.00	5.00
Process	Upper	23	4.00	.61134	.12747	3.7356	4.2644	2.67	5.00
	Total	43	3.97	.58560	.08930	3.7888	4.1492	2.67	5.00
	Lower	14	3.95	.76749	.20512	3.5058	4.3921	2.00	4.71
Outcomes	Upper	9	3.75	.70991	.23664	3.2003	4.2917	2.71	4.57
	Total	23	3.87	.73592	.15345	3.5513	4.1878	2.00	4.71
	Lower	16	3.92	.62657	.15664	3.5828	4.2505	2.67	4.67
Satisfaction	Upper	15	3.71	.75453	.19482	3.2933	4.1290	2.00	4.67
	Total	31	3.82	.68766	.12351	3.5650	4.0694	2.00	4.67
	Lower	11	3.34	.88933	.26814	3.7434	3.9384	1.75	4.50
Input	Upper	6	3.21	1.08877	.44449	3.0657	4.3509	1.75	5.00
	Total	17	3.29	.93222	.22610	3.8148	3.7734	1.75	5.00

Descriptive statistics for individual data items effectiveness ratings by principals in the upper and lower groups are presented in Table 6. Individual data items are listed from highest mean effectiveness rating to the lowest mean effectiveness rating of the upper performance group of principals. Mean ratings for 16 of 18 individual data items were above the 3.0 mid-point on the effectiveness rating scale for both groups indicating that both groups perceive the individual data items as being moderately to very effective in their efforts to increase mathematics scores on standardized assessments like the PSSA.

As expected, data collected from local benchmark assessments and standardized score data were rated as most effective with data collected from local benchmark assessments being rated as slightly more effective by both groups. Though principals in both groups rated these two individual outcomes type data items slightly higher than all others, principals in the upper performance group rated process type data items as more effective overall in their efforts to increase mathematics scores. Principals in the upper group rated "data about curriculum needs" as the third most effective individual data item. The effectiveness rating for this individual type of process data suggests that principals in the upper performance group believe that data about curriculum issues related to the teaching of mathematics is valuable in their efforts to increase mathematics scores. Though it was rated as the fourth most effective individual data item, principals in the lower group appear to believe strongly in the use of data about curriculum needs also as a key element of process data to increasing mathematics scores on standardized tests. This individual data item was the highest rated process type data item in both groups.

It should be noted that principals in the upper schools rated "teacher generated authentic assessment data" as the fourth most effective individual data item overall and the third most effective individual data item under the outcomes data type sub-category. This item was

Table 6

Principals' Effectiveness Rating for Individual Data Types

	Data Type	School	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Min	Max
Data collected from local benchmark assessments	outcomes	Lower	74	4.26	1.00765	.11714	4.0233	4.4902	1.00	5.00
		Upper	55	4.27	.78066	.10526	4.0617	4.4838	2.00	5.00
		Total	129	4.26	.91446	.08051	4.1043	4.4229	1.00	5.00
Standardized test score data	outcomes	Lower	61	4.05	.84511	.10821	3.8327	4.2656	1.00	5.00
		Upper	57	4.12	.68322	.09049	3.9415	4.3041	2.00	5.00
		Total	118	4.08	.76880	.07077	3.9446	4.2249	1.00	5.00
Data about curriculum needs	process	Lower	45	3.98	.83907	.12508	3.7257	4.2299	2.00	5.00
		Upper	51	4.08	.62748	.08786	3.9020	4.2549	3.00	5.00
		Total	96	4.03	.73202	.07471	3.8829	4.1796	2.00	5.00
Teacher generated authentic assessment data	outcomes	Lower	40	3.68	.97106	.15354	3.3644	3.9856	1.00	5.00
		Upper	39	3.90	.82062	.13140	3.6314	4.1634	2.00	5.00
		Total	79	3.78	.90115	.10139	3.5830	3.9867	1.00	5.00
Data about best practices for instruction	process	Lower	39	4.08	.66430	.10637	3.8616	4.2923	3.00	5.00
		Upper	38	3.89	.72743	.11801	3.6556	4.1338	2.00	5.00
		Total	77	3.99	.69762	.07950	3.8287	4.1454	2.00	5.00
Data about school program effectiveness	process	Lower	43	3.60	.95468	.14559	3.3108	3.8985	1.00	5.00
		Upper	43	3.77	.71837	.10955	3.5464	3.9885	2.00	5.00
		Total	86	3.69	.84382	.09099	3.5051	3.8670	1.00	5.00
Feedback data from teachers	satisfaction	Lower	56	3.55	.95193	.12721	3.2986	3.8085	2.00	5.00
		Upper	51	3.73	.91823	.12858	3.4672	3.9837	2.00	5.00
		Total	107	3.64	.93559	.09045	3.4562	3.8148	2.00	5.00
Value added assessment data	outcomes	Lower	37	3.84	1.09325	.17973	3.4733	4.2023	1.00	5.00
		Upper	33	3.67	.73598	.12812	3.4057	3.9276	2.00	5.00
		Total	70	3.76	.93925	.11226	3.5332	3.9811	1.00	5.00
Student attendance data	outcomes	Lower	43	3.72	1.16139	.17711	3.3635	4.0784	1.00	5.00
		Upper	35	3.63	.87735	.14830	3.3272	3.9300	2.00	5.00
		Total	78	3.68	1.03815	.11755	3.4454	3.9136	1.00	5.00

Table 6 (continued)

Principals' Effectiveness Rating for Individual Data Types

	Data Type	School	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Min	Max
Feedback data from students	satisfaction	Lower	33	3.52	.93946	.16354	3.1820	3.8483	1.00	5.00
		Upper	26	3.62	.85215	.16712	3.2712	3.9596	2.00	5.00
		Total	59	3.56	.89580	.11662	3.3259	3.7928	1.00	5.00
Student language proficiency data	input	Lower	16	3.56	.89209	.22302	3.0871	4.0379	2.00	5.00
		Upper	13	3.54	1.05003	.29123	2.9039	4.1730	2.00	5.00
		Total	29	3.55	.94816	.17607	3.1911	3.9124	2.00	5.00
Report card grade data	outcomes	Lower	44	3.48	.84876	.12796	3.2192	3.7353	2.00	5.00
		Upper	50	3.44	.76024	.10751	3.2239	3.6561	1.00	5.00
		Total	94	3.46	.79872	.08238	3.2939	3.6210	1.00	5.00
Teacher observation data	outcomes	Lower	50	3.56	.97227	.13750	3.2837	3.8363	1.00	5.00
		Upper	48	3.40	.86884	.12541	3.1435	3.6481	1.00	5.00
		Total	98	3.48	.92201	.09314	3.2947	3.6644	1.00	5.00
Student socioeconomic data	input	Lower	40	3.50	1.13228	.17903	3.1379	3.8621	1.00	5.00
		Upper	30	3.37	.80872	.14765	3.0647	3.6686	2.00	5.00
		Total	70	3.44	1.00196	.11976	3.2039	3.6818	1.00	5.00
Data on perceptions of the learning environment	satisfaction	Lower	21	3.76	.83095	.18133	3.3837	4.1401	2.00	5.00
		Upper	22	3.36	1.00216	.21366	2.9193	3.8080	1.00	5.00
		Total	43	3.56	.93356	.14237	3.2708	3.8454	1.00	5.00
Student ethnicity data	input	Lower	18	3.17	1.29479	.30518	2.5228	3.8105	1.00	5.00
		Upper	21	2.76	1.26114	.27520	2.1878	3.3360	1.00	5.00
		Total	39	2.95	1.27628	.20437	2.5350	3.3624	1.00	5.00
Student gender data	input	Lower	43	2.21	1.03643	.15805	1.8903	2.5283	1.00	5.00
		Upper	35	2.29	.98731	.16689	1.9466	2.6249	1.00	5.00
		Total	78	2.24	1.00887	.11423	2.0161	2.4711	1.00	5.00

followed closely by "data about best practices for instruction" and "data about school program effectiveness." Though "feedback data from teachers" and "teacher observation data" were reported as used by a high percentage of principals in both groups, principals' rating of effectiveness for these individual data types was comparatively lower. However, principals in the upper group rated feedback data from teachers as the most effective individual data item under the satisfaction data type sub-category.

Principals in both performance groups rated student language proficiency as the most effective individual data item under the input data type sub-category. Socioeconomic data was rated as the next most effective individual data item by both groups also. However, the mean rating of effectiveness by principals in the lower group for this individual data item was slightly higher. The mean ratings for both groups had "student ethnicity data" as the next highest rated input type data. Again, principals in the lower group rated this individual data item slightly higher than principals in the upper group. Both groups reported the lowest effectiveness rating in this sub-category for "student gender data."

Use of Data Tools

The next research question addressed by the study sought to determine data tools secondary principals in high and low performing schools use and perceive as effective for increasing student mathematics performance on standardized assessments. Principals responded either yes or no to items that asked if they used any of nine specific data tools in their efforts to increase mathematics scores on standardized assessments. Yes responses were recorded numerically as a 2 and no responses were recorded numerically as a 1 in order to generate overall means for comparison of principals' use of data tools. Descriptive statistics on principals' use of data tools are presented in Table 7. Data tools are ordered from the highest mean to the lowest mean of principals in the upper schools indicating use.

Not surprisingly, the highest mean responses by principals in both groups were found for data tools that are designed to analyze outcomes type data. These tools included the PAAYP data tool, the Emetric data tool, the PVAAS data tool, and the 4Sight data tool. Specifically, the PAAYP data tool and the Emetric data tool are designed to work with standardized test score data, while the 4Sight data tool is designed to work with data collected from locally given benchmark assessments. Overall, these data tools were cited as used by the most principals among nine individual data tools. A higher percentage of principals in the lower group reported use of these three data tools with over 90% indicating use of the PAAYP data tool and the 4Sight data tool.

A surprising number of principals in both groups reported use of the Pennsylvania Value Added System data tool. In the lower group, 79% of principals reported use of this tool while 69% of principals in the upper group indicated use. These percentages are high considering that less than 50% of principals in both groups cited the use of value added data in efforts to increase student achievement in mathematics on standardized assessments.

Slightly more principals in the upper group reported the use of Excel spreadsheets in efforts to increase their mathematics scores on standardized assessments. A total of 56% of principals in both groups reported the use of a software based student management system as a data tool. Few principals in both groups reported the use of a school or district developed data tool. The NAEP data tool was used by the lowest number of principals in both groups with only 13% of principals in the upper group reporting use and 9% in the lower group citing use.

Table 7

Principals' Use of Data Tools

Tool	School	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Min	Max
PAAYP	Lower	75	1.93	.251	.029	1.88	1.99	1	2
	Upper	72	1.85	.362	.043	1.76	1.93	1	2
	Total	147	1.89	.313	.026	1.84	1.94	1	2
Emetric	Lower	75	1.83	.381	.044	1.74	1.91	1	2
	Upper	72	1.81	.399	.047	1.71	1.90	1	2
	Total	147	1.82	.389	.032	1.75	1.88	1	2
Pennsylvania Value Added System	Lower	75	1.79	.412	.048	1.69	1.88	1	2
	Upper	72	1.69	.464	.055	1.59	1.80	1	2
	Total	147	1.74	.439	.036	1.67	1.81	1	2
4Sight	Lower	75	1.91	.293	.034	1.84	1.97	1	2
	Upper	72	1.65	.479	.057	1.54	1.77	1	2
	Total	147	1.78	.414	.034	1.71	1.85	1	2
Student Management System	Lower	75	1.56	.500	.058	1.45	1.67	1	2
	Upper	72	1.56	.500	.059	1.44	1.67	1	2
	Total	147	1.56	.498	.041	1.48	1.64	1	2
Excel Spreadsheet	Lower	75	1.44	.500	.058	1.33	1.55	1	2
	Upper	72	1.51	.503	.059	1.40	1.63	1	2
	Total	147	1.48	.501	.041	1.39	1.56	1	2
School Developed	Lower	75	1.27	.445	.051	1.16	1.37	1	2
	Upper	72	1.19	.399	.047	1.10	1.29	1	2
	Total	147	1.23	.423	.035	1.16	1.30	1	2
District Developed	Lower	75	1.13	.342	.040	1.05	1.21	1	2
	Upper	72	1.17	.375	.044	1.08	1.25	1	2
	Total	147	1.15	.358	.030	1.09	1.21	1	2
NAEP	Lower	75	1.09	.293	.034	1.03	1.16	1	2
	Upper	72	1.13	.333	.039	1.05	1.20	1	2
	Total	147	1.11	.313	.026	1.06	1.16	1	2

Principals in the upper and lower groups who indicated use of an individual tool were asked to rate the effectiveness of the data tool in their efforts to improve mathematics scores on standardized assessments. Principals were asked to rate the effectiveness of data tools on a five item scale from “not effective” to “very effective.” Each item on the scale was assigned a number in the analysis for the purpose of generating a mean effectiveness rating. The rating of “not effective” was given the numerical value of 1. The next rating, interpreted as “slightly effective”, was given the numerical value of 2. The next item, interpreted as “somewhat effective” on the scale was given the value equivalent of 3. The next item, interpreted as “effective”, was given the numerical value of 4. The final rating on the scale, “very effective,” was assigned the numerical value of 5.

Principals’ mean ratings of data tool effectiveness were calculated for each individual data tool. Results are presented in Table 8 in order of highest mean to lowest mean by the upper performance group. Principals rated the data tools on a five item scale from not effective to very effective. All data tools except one were rated above the midpoint of 3.0 on the effectiveness rating scale by both groups. The NAEP data tool had a mean rating of 2.75 by principals in the upper group, while principals in the lower group rated it higher with a 3.86 mean rating. The NAEP tool was the only data tool with a reported rating less than 3.0 for effectiveness.

It should be noted that though relatively few principals reported use of a district or school created data tool, principals in the upper group rated these tools as the most effective they use in their efforts to increase mathematics scores. Principals in the lower group rated the 4Sight data tool as the most effective tool. The 4Sight data tool had the third highest mean rating among principals in the upper schools following district and school developed data tools. The Emetric and PAAYP data tools that had high reported usage among principals in both groups were rated

Total 8

Principals' Rating of Individual Tool Effectiveness

Tool	School	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Min	Max
	Lower	10	3.80	.91894	.29059	3.1426	4.4574	2.00	5.00
District	Upper	11	4.18	.60302	.18182	3.7767	4.5869	3.00	5.00
Developed	Total	21	4.00	.77460	.16903	3.6474	4.3526	2.00	5.00
	Lower	20	4.05	.75915	.16975	3.6947	4.4053	3.00	5.00
School	Upper	14	4.14	.77033	.20588	3.6981	4.5876	3.00	5.00
Developed	Total	34	4.09	.75348	.12922	3.8253	4.3511	3.00	5.00
	Lower	68	4.10	1.08090	.13108	3.8413	4.3646	1.00	5.00
4Sight	Upper	47	3.89	.96084	.14015	3.6115	4.1757	1.00	5.00
	Total	115	4.02	1.03435	.09645	3.8263	4.2085	1.00	5.00
	Lower	33	3.97	.80951	.14092	3.6827	4.2567	2.00	5.00
Excel	Upper	35	3.83	.70651	.11942	3.5859	4.0713	3.00	5.00
Spreadsheet	Total	68	3.90	.75587	.09166	3.7141	4.0800	2.00	5.00
	Lower	62	3.58	.87868	.11159	3.3575	3.8038	2.00	5.00
Emetric	Upper	56	3.63	.70227	.09384	3.4369	3.8131	2.00	5.00
	Total	118	3.60	.79679	.07335	3.4564	3.7470	2.00	5.00
Pennsylvania	Lower	70	3.60	.98393	.11760	3.3654	3.8346	1.00	5.00
AYP	Upper	59	3.49	.83816	.10912	3.2731	3.7100	1.00	5.00
	Total	129	3.55	.91823	.08085	3.3904	3.7104	1.00	5.00
Student	Lower	42	3.60	.88509	.13657	3.3194	3.8711	2.00	5.00
Management	Upper	38	3.47	.64669	.10491	3.3261	3.6862	2.00	5.00
System	Total	80	3.54	.77857	.08705	3.3642	3.7108	2.00	5.00
Pennsylvania	Lower	59	3.39	1.03419	.13464	3.1203	3.6593	1.00	5.00
Value Added	Upper	49	3.39	.73076	.10439	3.1779	3.5977	2.00	5.00
System	Total	108	3.39	.90516	.08710	3.2162	3.5616	1.00	5.00
	Lower	7	3.86	1.46385	.55328	2.5033	5.2110	1.00	5.00
NAEP	Upper	8	2.75	.88641	.31339	2.0089	3.4911	1.00	4.00
	Total	15	3.27	1.27988	.33046	2.5579	3.9754	1.00	5.00

slightly less effective than the 4Sight data tool and Excel spreadsheets by both groups. Mean ratings for the Pennsylvania Value Added System data tool were the same in both groups. This data tool was rated only more effective than the NAEP tool.

Use of Data Systems

The next question in the research study was focused on determining the extent that secondary principals in high and low performing schools use data systems to increase student achievement in mathematics on standardized assessments. Respondents were asked to rate for 13 items the extent that they use elements of data systems on a five item scale that ranged from "to no extent" at lowest end to the highest end, "to a great extent." Each item on the scale was assigned a number in the analysis for the purpose of generating a mean extent of use rating. The rating, "to no extent," was given the numerical value of 1. The next rating, interpreted as "to a slight extent," was given the numerical value of 2. The next item, interpreted as "to some extent" on the scale was given the numerical value of 3. The next item, interpreted as "to a moderate extent," was given the numerical value of 4. The final rating on the scale, "to a great extent," was assigned the numerical value of 5.

The13 items were grouped together based around a common content focus to form the sub-category use of data systems. Descriptive statistics for principals' use of data systems are presented in Table 9. The upper group had a mean of 3.79 for their reported use of data systems, while the lower group reported a mean of 3.95. It appears that principals in the lower group perceive that they use data systems to a greater extent in their efforts to increase student mathematics scores on standardized assessments than principals in the upper group.

Table 9

Principals' Use of Data Systems

	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Min	Max
Lower	75	3.95	.63976	.07387	3.8015	4.0959	2.54	5.00
Upper	72	3.79	.58186	.06857	3.6496	3.9231	2.38	4.85
Total	147	3.87	.61541	.05076	3.7689	3.9695	2.38	5.00

Analysis of Differences

The final research question in the study had five sub-questions designed to guide analysis of differences between secondary principals in high and low performing schools in their use of data driven decision making to enhance student achievement in mathematics on standardized assessments. These sub-questions sought to determine if a difference existed between secondary principals in high and low performing schools in their use of data types, perception of data type effectiveness, use of data tools, perception of data tool effectiveness, and the extent that they use data systems. One-way ANOVAs were used to determine if significant differences existed between the upper schools and lower schools on use of input data, use of process data, use of outcomes data, use of satisfaction data, perceived effectiveness of input data, perceived effectiveness of process data, perceived effectiveness of outcomes data, perceived effectiveness of satisfaction data, use of data systems, use of data tools, and perceived effectiveness of data tools used (sub-categories 1-11). Chi-square analysis was used to provide further analysis of differences between principals' use of input data, use of outcomes data, use of process data, use of satisfaction data, and use of data tools. A two-way ANOVA was conducted to examine differences in principal responses for all sub-categories by performance group and principal age. Additional two-way ANOVAs were completed to analyze difference by performance group and principal years experience in the school and by performance group and district size for all sub-categories.

The first one-way ANOVA sought to determine if a difference existed between principals in the high and low performing schools regarding their use of different data types. Principals were asked to respond either yes or no to indicate if they used an individual data item in their efforts to increase mathematics scores. Responses of yes were recorded as a 2 and responses of no were recorded as a 1 in order to generate a mean use for each data type including input, outcomes, process, and satisfaction data. The results of the ANOVA presented in Table 10 show no statistically significant difference between principals' in low and high performing schools use of input data, $F(1, 145) = .441$, $p > .05$, use of process data, $F(1, 145) = .725$, $p > .05$, use of outcomes data, $F(1, 145) = .717$, $p > .05$, or use of satisfaction data, $F(1, 145) = .281$, $p > .05$. Although there were no statistically significant differences, slightly higher means are found in three of four sub-categories for principals in the lower group (see Table 3). Principals in the upper group have a slightly higher mean for the use of process data.

Table 10

One-Way ANOVA for Performance and Use of Data Type

Data Type		Sum of Squares	df	Mean Square	F	Sig.
Input	Between Groups	.766	1	.766	.441	.508
	Within Groups	251.778	145	1.736		
	Total	252.544	146			
Outcomes	Between Groups	2.306	1	2.306	.717	.398
	Within Groups	466.306	145	3.216		
	Total	468.612	146			
Process	Between Groups	.720	1	.720	.725	.396
	Within Groups	143.947	145	.993		
	Total	144.667	146			
Satisfaction	Between Groups	.309	1	.309	.281	.597
	Within Groups	159.542	145	1.100		
	Total	159.850	146			

Note. $p > .05$.

Chi-Square Analysis of Data Type Use

A Chi-Square analysis was completed on survey items included in the sub-categories use of input data, use of outcomes data, use of process data, and use of satisfaction data to analyze mean differences between principals in the upper and lower performance groups for individual data items. Though 16 of 17 items showed no significant difference between the groups, one item within the sub-category use of outcomes data showed a significant difference between performance groups (see Table 11). The following narrative analyzes how the groups differed.

Table 11 presents a chi-square analysis of all items that asked respondents to identify if they used an individual data item. There was a significant difference found between the upper and lower schools in the use of data collected from local benchmark assessments, χ^2 (1, N = 147) = 16.967, p = .000). The data in Table 11 shows that more principals in lower performing schools use data collected from local benchmark assessments than principals in higher schools in their efforts to improve their schools' mathematics scores on standardized tests.

A second one-way ANOVA sought to determine differences between principals in high and low performing schools in their mean effectiveness rating for different data types. Results presented in Table 12 show no significant differences were found between principals in the upper and lower groups in their effectiveness rating of the data types. However, slightly higher mean effectiveness ratings were reported by principals in the low performing schools for outcomes, input, and satisfaction data types. Principals in the upper group reported a slightly higher mean effectiveness rating for process type data. Descriptive statistics for this analysis are shown in Table 5.

Table 11

Chi-Square for Use of Individual Data Items and Performance

Individual Data Item			Performance lower	upper	Total	df	sig.
student ethnicity data	No	Count	57.0	51.0	108.0		
		Expected Count	55.1	52.9	108.0		
	Yes	Count	18.0	21.0	39.0		
		Expected Count	19.9	19.1	39.0		
	Total	Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	0.478
student language proficiency data	No	Count	59.0	59.0	118.0		
		Expected Count	60.2	57.8	118.0		
	Yes	Count	16.0	13.0	29.0		
		Expected Count	14.8	14.2	29.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	0.618
feedback data from teachers	No	Count	19.0	21.0	40.0		
		Expected Count	20.0	20.0	40.0		
	Yes	Count	56.0	51.0	107.0		
		Expected Count	55.0	52.0	107.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	0.602
teacher observation data	No	Count	25.0	24.0	49.0		
		Expected Count	25.0	24.0	49.0		
	Yes	Count	50.0	48.0	98.0		
		Expected Count	50.0	48.0	98.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	1.000
student socioeconomic data	No	Count	35.0	42.0	77.0		
		Expected Count	39.0	38.0	77.0		
	Yes	Count	40.0	30.0	70.0		
		Expected Count	36.0	34.0	70.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		

	Pearson Chi-square					1	0.157
report card grade data	No	Count	31.0	22.0	53.0		
		Expected Count	27.0	26.0	53.0		
	Yes	Count	44.0	50.0	94.0		
		Expected Count	48.0	46.0	94.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	0.174
teacher generated authentic assessment data	No	Count	35.0	33.0	68.0		
		Expected Count	35.0	33.0	68.0		
	Yes	Count	40.0	39.0	79.0		
		Expected Count	40.0	39.0	79.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	0.919
data collected from local benchmark assessments	No	Count	1.0	17.0	18.0		
		Expected Count	9.0	9.0	18.0		
	Yes	Count	74.0	55.0	129.0		
		Expected Count	66.0	63.0	129.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	0.000
data perception of the learning environment	No	Count	54.0	50.0	104.0		
		Expected Count	53.0	51.0	104.0		
	Yes	Count	21.0	22.0	43.0		
		Expected Count	22.0	21.0	43.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	0.733
feedback data from students	No	Count	42.0	46.0	88.0		
		Expected Count	45.0	43.0	88.0		
	Yes	Count	33.0	26.0	59.0		
		Expected Count	30.0	29.0	59.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	0.329

data about school program effectiveness	No	Count	32.0	29.0	61.0		
		Expected Count	31.0	30.0	61.0		
	Yes	Count	43.0	43.0	86.0		
		Expected Count	44.0	42.0	86.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	0.769
value added assessment data	No	Count	38.0	39.0	77.0		
		Expected Count	39.0	38.0	77.0		
	Yes	Count	37.0	33.0	70.0		
		Expected Count	36.0	34.0	70.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	0.671
standardized test score data	No	Count	14.0	15.0	29.0		
		Expected Count	15.0	14.0	29.0		
	Yes	Count	61.0	57.0	118.0		
		Expected Count	60.0	58.0	118.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	0.741
student gender data	No	Count	59.0	60.0	119.0		
		Expected Count	61.0	58.0	119.0		
	Yes	Count	16.0	12.0	28.0		
		Expected Count	14.0	14.0	28.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	0.471
data about curriculum needs	No	Count	30.0	21.0	51.0		
		Expected Count	26.0	25.0	51.0		
	Yes	Count	45.0	51.0	96.0		
		Expected Count	49.0	47.0	96.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	0.168

data about best practices for instruction	No	Count	36.0	34.0	70.0		
		Expected Count	36.0	34.0	70.0		
	Yes	Count	39.0	38.0	77.0		
		Expected Count	39.0	38.0	77.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	0.925
student attendance data	No	Count	32.0	37.0	69.0		
		Expected Count	35.0	34.0	69.0		
	Yes	Count	43.0	35.0	78.0		
		Expected Count	40.0	38.0	78.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	0.289

Table 12

One-Way ANOVA for Performance Principals' Effectiveness Rating of Data Types

		Sum of Squares	df	Mean Square	F	Sig.
Rating of Input Data	Between Groups	.068	1	.068	.074	.789
	Within Groups	13.836	15	.922		
	Total	13.904	16			
Rating of Outcomes Data	Between Groups	.226	1	.226	.405	.531
	Within Groups	11.689	21	.557		
	Total	11.915	22			
Rating of Process Data	Between Groups	.048	1	.048	.136	.714
	Within Groups	14.356	41	.350		
	Total	14.403	42			
Rating of Satisfaction Data	Between Groups	.327	1	.327	.684	.415
	Within Groups	13.859	29	.478		
	Total	14.186	30			

Note. $p > .05$.

The researcher next conducted a one-way ANOVA to determine if a difference existed in use of data tools between principals in high and low performing schools. Principals were asked to identify if they used a specific data tool. Yes responses were assigned a numerical value of 2 and no responses were assigned a numerical value of 1 in order to generate means for overall data tool use (Table 13). The results presented in Table 14 reveal no significant difference between principals in high and low performing schools in their overall use of data tools, $F(1, 145) = 2.241$, $p > .05$. Descriptive statistics for overall reported use of data tools are shown in Table 13. Principals in the lower group show a slightly higher mean for the use of data tools than principals in the upper group.

Table 13

Principals' Overall Use of Data Tools

	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Min	Max
Lower	75	1.55	.16530	.01909	1.5116	1.5877	1.00	2.00
Upper	72	1.51	.18638	.02197	1.4624	1.5500	1.00	1.89
Total	147	1.53	.17669	.01457	1.4995	1.5571	1.00	2.00

Table 14

One-Way ANOVA for Performance and Principals' Overall Use of Data Tools

	Sum of Squares	df	Square	F	Sig.
Between Groups	.069	1	.069	2.241	.137
Within Groups	4.488	145	.031		
Total	4.558	146			

Note. p>.05.

Chi-Square Analysis of Data Tool Use

A Chi-Square analysis was completed on survey items included in the sub-category use of data tools to analyze mean differences between principals in the upper and lower performance groups. The results of this analysis are presented in Table 15. One item that asked principals if they used “4Sight Assessments/Success For All Foundation data tool” in efforts to increase mathematics scores on standardized assessments within the sub-category use of data tools showed a significant difference between the two performance groups, χ^2 (1, N = 147) = 13.904, *p* = .000). Results indicate that more principals in the lower performing schools use the 4Sight Assessments/Success For All Foundation data tool more than principals in the higher performing schools in efforts to improve student achievement in mathematics.

The researcher next analyzed differences between principals in each group for their overall effectiveness rating of data tools. A one-way ANOVA was conducted to analyze principals’ means for the overall effectiveness rating of data tools (sub-category 11). Descriptive statistics for this analysis are shown in Table 16. Principals were asked to rate data tools for effectiveness on a five item scale ranging from not effective to very effective. Table 16 reflects a smaller sample size from each group to generate the overall effectiveness rating for data tools as means were calculated from principals that indicated use of all data tools. The results of this analysis, F (1, 19) = .104, p >.05, indicate no significant difference between the two groups and are presented in Table 17. Though no significant difference was found between groups for principals’ overall effectiveness rating of data tools, a slightly higher mean is found for principals in the lower group (see Table 16).

Table 15

Chi-Square for Use of Individual Data Tools and Performance

Individual Data Tool			Performance lower	Performance upper	Total	df	sig.
PAAYP	No	Count	5.0	11.0	16.0		
		Expected Count	8.2	7.8	16.0		
	Yes	Count	70.0	61.0	131.0		
		Expected Count	66.8	64.2	131.0		
	Total	Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	.094
Emetric	No	Count	13.0	14.0	27.0		
		Expected Count	13.8	13.2	27.0		
	Yes	Count	62.0	58.0	120.0		
		Expected Count	61.2	58.8	120.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	.741
PA Value Added System	No	Count	16.0	22.0	38.0		
		Expected Count	19.4	18.6	38.0		
	Yes	Count	59.0	50.0	109.0		
		Expected Count	55.6	53.4	109.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	.202
4Sight Assessments/ Success For All Foundation	No	Count	7.0	25.0	32.0		
		Expected Count	16.3	15.7	32.0		
	Yes	Count	68.0	47.0	115		
		Expected Count	58.7	56.3	115.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	.000
Student Mngmt System	No	Count	33.0	32.0	65.0		
		Expected Count	33.2	31.8	65.0		
	Yes	Count	42.0	40.0	82.0		
		Expected Count	41.8	40.2	82.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		

	Pearson Chi-square					1	.957
Excel Spreadsheet	No	Count	42.0	35.0	77.0		
		Expected Count	39.3	37.7	77.0		
	Yes	Count	33.0	37.0	70.0		
		Expected Count	35.7	34.3	70.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	.370
School Developed	No	Count	55.0	58.0	113.0		
		Expected Count	57.7	55.3	113.0		
	Yes	Count	20.0	14.0	34.0		
		Expected Count	17.3	16.7	34.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	.299
District Developed	No	Count	65.0	60.0	125.0		
		Expected Count	63.8	61.2	125.0		
	Yes	Count	10.0	12.0	22.0		
		Expected Count	11.2	10.8	22.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	.571
NAEP	No	Count	68.0	63.0	131.0		
		Expected Count	66.8	64.2	131.0		
	Yes	Count	7.0	9.0	16.0		
		Expected Count	8.2	7.8	16.0		
	Total	Count	75.0	72.0	147.0		
		Expected Count	75.0	72.0	147.0		
	Pearson Chi-square					1	.538

Table 16

Principals' Overall Data Tool Effectiveness Rating

	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Min	Max
Lower	10	3.90	.54638	.17278	3.5089	4.2907	3.00	4.78
Upper	11	3.83	.46888	.14137	3.5134	4.1434	2.83	4.50
Total	21	3.86	.49558	.10815	3.6368	4.0880	2.83	4.78

Table 17

One-Way ANOVA for Performance and Overall Data Tool Effectiveness Rating

	Sum of Squares	df	Square	F	Sig.
Between Groups	.027	1	.027	.104	.751
Within Groups	4.885	19	.257		
Total	4.912	20			

Note. $p>.05$.

A final one-way ANOVA was conducted to exam any differences between the principals in the upper and lower groups in their use of data systems. Principals were asked to identify the extent that they used an element of data systems to increase student achievement on standardized mathematics assessments on a five item scale ranging from “to no extent” through “to a great extent.” Table 18 shows the results of this analysis. A significant difference was not found between the two groups for their use of data systems, $F(1, 145) = 2.586$, $p > .05$. However, descriptive statistics show principals in the low performing schools report a higher mean use of data systems (see Table 9).

Table 18

One-Way ANOVA for Performance and Use of Data Systems

	Sum of Squares	df	Square	F	Sig.
Between Groups	.969	1	.969	2.586	.110
Within Groups	54.326	145	.375		
Total	55.295	146			

Note. p>.05.

Two-Way ANOVA Analysis

A series of two-way ANOVAs were conducted to examine differences in means for all sub-categories. This two factor analysis used the main factor of performance with the factors of age, district size, and years experience as principals at current schools. An additional two factor analysis was performed using the main factors of gender and performance. The two-way

ANOVA is designed to examine the effects of more than one factor simultaneously and any significance the effects may have through their interaction (Huck, 2004). Through the use of the two-way ANOVA, some of the random variability is explained by an additional factor so that significant differences can be more easily revealed. The main focus of this study was to examine differences between the performance groups. Analyzing the data using main factors in addition to the main factor of performance was beneficial to finding any significant differences between the groups. The two-way ANOVA also examines the interaction of two main factors so that it can be determined whether or not the effect of one variable changes with the level of the other factor. These analyses necessarily include independent variables not specifically stated in the research questions. These findings were, therefore, presented separately from the preceding analyses. The analyses aided in finding significant differences between the upper and lower performance groups. Additional significant findings from these analyses related to age, school size, years experience as principal at current school, and gender are also presented.

A Look at the Impact of Age and Performance

A two-way ANOVA was conducted by the researcher to determine if significant differences existed in mean scores for perceived use of data systems by age and performance. This two-way ANOVA revealed significant differences in the main effects of performance and age (see Tables 19-20). Descriptive statistics are presented in Table 19. The main effect of performance yielded an F ratio of $F(1, 146) = 4.112$, $p < .05$, indicating that the mean score for lower schools ($M = 3.94$, $SD = .63$) differed significantly from that of upper schools ($M = 3.78$, $SD = .58$). Surprisingly, the data presented in Table 20 suggests that principals in the higher performing schools actually use data systems to a lesser extent than principals in the lower performing schools. The main effect of age yielded an F ratio of $F(2, 146) = 3.738$, $p < .05$,

Table 19

Descriptive Statistics for Performance/Age and Perceived Use of Data Systems

Performance	Age	Mean	Std. Deviation	N
Lower	25-39 years	3.93	.76	22
	40-54 years	3.85	.57	41
	55+ years	4.31	.49	12
	Total	3.94	.63	75
Upper	25-39 years	3.66	.54	19
	40-54 years	3.76	.63	38
	55+ years	4.00	.42	15
	Total	3.78	.58	72
Total	25-39 years	3.80	.67	41
	40-54 years	3.80	.60	79
	55+ years	4.13	.47	27
	Total	3.86	.61	147

Table 20

Two-Way ANOVA for Performance/Age and Perceived Use of Data Systems

Source	Type III Sum of Squares	*df*	Mean Square	*F*	Sig.
Corrected Model	3.953(a)	5	.791	2.171	.061
Intercept	1851.907	1	1851.907	5085.931	.000
Performance	1.497	1	1.497	4.112	.044
Age	2.723	2	1.361	3.738	.026
Performance* Age	.360	2	.180	.494	.611
Error	51.341	141	.364		
Total	2255.964	147			
Corrected Total	55.295	146			

Note. a R Squared = .071 (Adjusted R Squared = .039).

indicating that the means for the age groups differed significantly. Post hoc analyses using the Tukey HSD revealed that the mean perceived use of data systems was significantly lower in the 40-54 year olds ($M = 3.80$, $SD = .60$) than in 55+ years old ($M = 4.13$, $SD = .47$). Though a mathematical significant difference was found, the closeness of the means suggests there may not be an educationally significant difference between the groups. The interaction effect was non-significant, $F(1, 146) = .494$, $p > .05$ (see Table 16).

The two-way ANOVA conducted by the researcher to analyze the sub-category scores of principals in the upper schools, lower schools, and principal age showed further significant differences among means. A significant difference was found for age groups and the use of satisfaction data. Descriptive statistics can be found in Table 21. The main effect of performance yielded an *F* ratio of $F(1, 146) = 1.199$, $p > .05$, indicating that the mean score for lower schools was similar to the mean for upper schools. The main effect of age yielded an F ratio of $F(2, 146) = 4.087$, $p < .05$, indicating that the mean for 25-39 year olds ($M = 4.51$, $SD = 1.05$), 40-54 year olds ($M = 4.22$, $SD = 1.01$) and 55+ years ($M = 4.85$, $SD = 1.02$) differed significantly. More specifically, post hoc analyses using the Tukey HSD indicated that the average use of satisfaction data was significantly lower in the 40-54 year olds ($M = 4.22$, $SD = 1.01$) than 55+ years old ($M = 4.85$, $SD = 1.02$) suggesting that principals in the 40-54 year old range use satisfaction data less in efforts to improve mathematics scores on standardized assessments. The interaction effect was non-significant, $F(1, 146) = .710$, $p > .05$ (see Table 22).

Table 21

Descriptive Statistics for Performance/Age and of Use Satisfaction Data

Performance	Age	Mean	Std. Deviation	N
Lower	25-39 years	4.63	.95	22
	40-54 years	4.19	1.03	41
	55+ years	5.08	1.08	12
	Total	4.46	1.05	75
Upper	25-39 years	4.36	1.16	19
	40-54 years	4.26	1.00	38
	55+ years	4.66	.97	15
	Total	4.37	1.04	72
Total	25-39 years	4.51	1.05	41
	40-54 years	4.22	1.01	79
	55+ years	4.85	1.02	27
	Total	4.42	1.04	147

Table 22

Two-Way ANOVA for Performance/Age and of Use Satisfaction Data

Source	Type III Sum of Squares	*df*	Mean Square	*F*	Sig.
Corrected Model	10.281(a)	5	2.056	1.938	.092
Intercept	2478.482	1	2478.482	2336.48	.000
Performance	1.272	1	1.272	1.199	.275
Age	8.670	2	4.335	4.087	.019
Performance* Age	1.506	2	.753	.710	.494
Error	149.569	141	1.061		
Total	3034.000	147			
Corrected Total	159.850	146			

Note. a R Squared = .064 (Adjusted R Squared = .031).

Further two-way ANOVA analysis by principal age and performance revealed a significant interaction effect between mean scores for perceived effectiveness ratings of input data (see Tables 23-24). Descriptive statistics can be found in Table 23. The interaction effect was significant, $F(1, 16) = 6.26$, $p < .05$ (see Table 24). The nature of this interaction is visible in Figure 4 (the means for 25-29 age group was not estimated due to small sample size). Those 55+ years old in the lower group ($M = 4.31$, $SD = .23$) had a significantly higher mean for effectiveness ratings of input data in increasing mathematics scores on standardized tests than 40-54 years old in lower schools ($M = 2.62$, $SD = .59$).

Table 23

Descriptive Statistics for Performance/Age and Perceived Effectiveness of Input Data

Performance	Age	Mean	Std. Deviation	N
Lower	25-39 years	3.00	.50	3
	40-54 years	2.62	.59	4
	55+ years	4.31	.23	4
	Total	3.34	.88	11
Upper	40-54 years	3.37	1.33	4
	55+ years	2.87	.53	4
	Total	3.20	1.08	8
Total	25-39 years	3.00	.50	3
	40-54 years	3.00	1.03	8
	55+ years	3.83	.80	8
	Total	3.29	.93	19

Table 24

Two-Way ANOVA for Performance/Age and Perceived Effectiveness of Input Data

Source	Type III Sum of Squares	*df*	Mean Square	*F*	Sig.
Corrected Model	6.576(a)	4	1.644	2.692	.082
Intercept	142.561	1	142.561	233.448	.000
Performance	.378	1	.378	.619	.447
Age	1.608	2	.804	1.316	.304
Performance* Age	3.828	1	3.828	6.269	.028
Error	7.328	12	.611		
Total	198.375	17			
Corrected Total	13.904	16			

Note. a R Squared = .473 (Adjusted R Squared = .297).

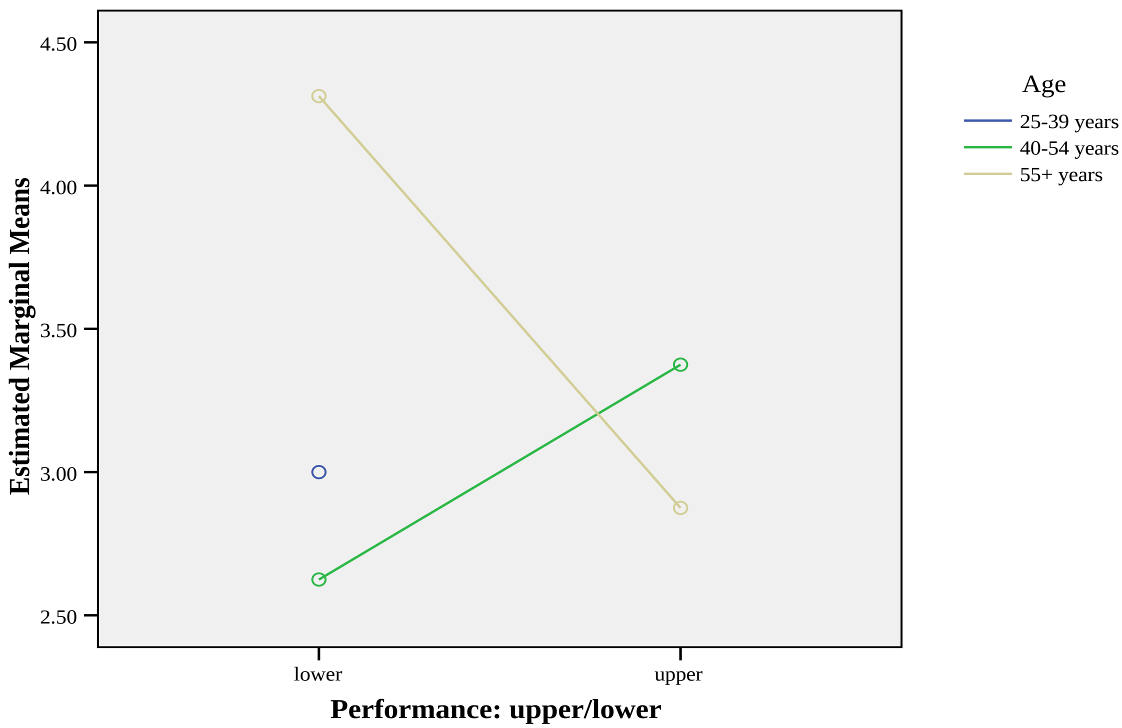

Figure 4.   Plot of age/performance and input data effectiveness.

An additional two-way ANOVA conducted to examine mean perceived effectiveness of process data by age and performance indicated a significant interaction effect. Descriptive statistics are presented in Table 25. The interaction effect was significant, $F(1, 42) = 4.11$, $p < .05$ (see Table 26). More specifically, principals in lower performing schools 55+ years old ($M = 4.88$, $SD = .19$) had significantly higher mean perceived effectiveness of process data scores than 40-54 year olds ($M = 3.70$, $SD = .45$) and 25-39 year olds ($M = 3.81$, $SD = .41$). See Figure 5.

Table 25

Descriptive Statistics for Performance/Age and Perceived Effectiveness of Process Data

Performance	Age	Mean	Std. Deviation	N
Lower	25-39 years	3.81	.41	9
	40-54 years	3.70	.45	8
	55+ years	4.88	.19	3
	Total	3.93	.56	20
Upper	25-39 years	3.91	.87	4
	40-54 years	4.05	.58	12
	55+ years	3.95	.59	7
	Total	4.00	.61	23
Total	25-39 years	3.84	.55	13
	40-54 years	3.91	.55	20
	55+ years	4.23	.66	10
	Total	3.96	.58	43

Table 26

Two-Way ANOVA for Performance/Age and Perceived Effectiveness of Process Data

Source	Type III Sum of Squares	*df*	Mean Square	*F*	Sig.
Corrected Model	3.399(a)	5	.680	2.286	.066
Intercept	566.423	1	566.423	1904.51	.000
Performance	.227	1	.227	7	.388
Age	1.952	2	.976	.764	.050
Performance* Age	2.449	2	1.225	3.282	.024
Error	11.004	37	.297	4.118	
Total	691.778	43			
Corrected Total	14.403	42			

Note. a R Squared = .236 (Adjusted R Squared = .133).

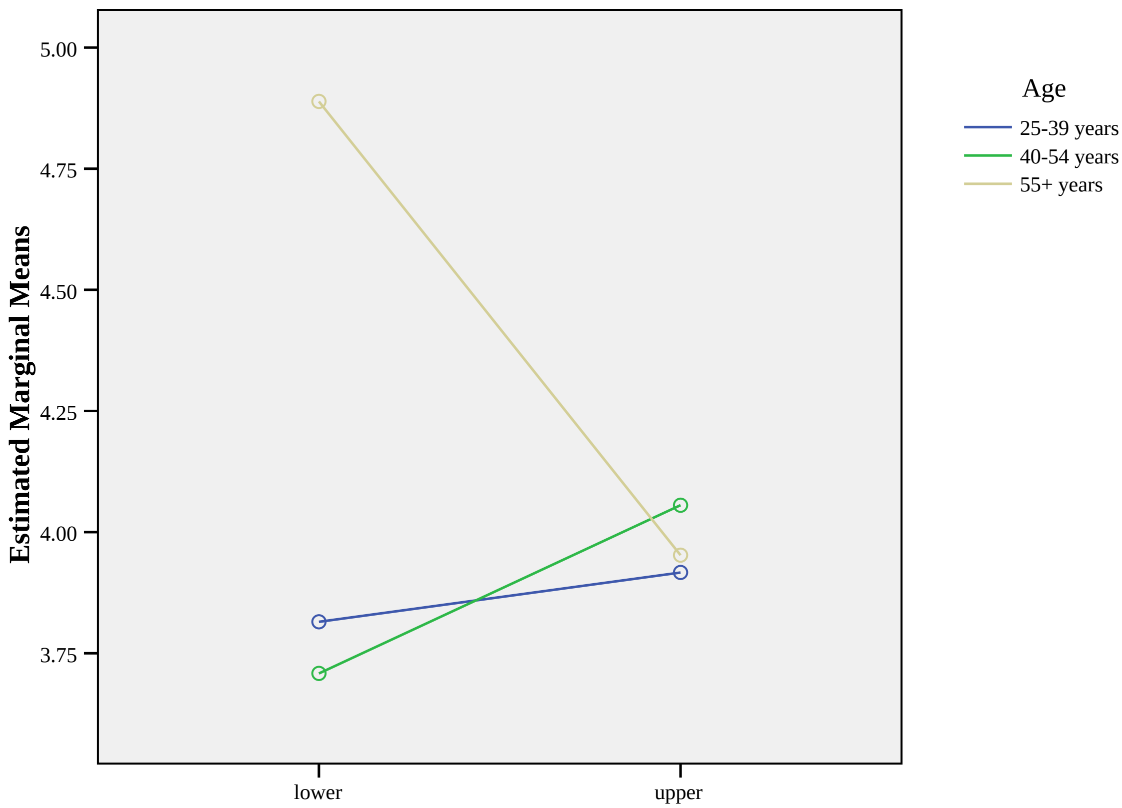

Figure 5. Plot of age/performance and process data effectiveness.

A Look at the Impact of Size and Performance

An additional two-way ANOVA was conducted by the researcher to analyze mean scores for all sub-categories to determine if significant differences existed between the upper schools, lower schools, and school size. A significant difference was found in the main effect of performance for use of data tools (see Tables 27-28) by principals in efforts to increase standardized mathematics assessment scores. Descriptive statistics are presented in Table 27. The main effect of performance yielded an F ratio of $F(1, 146) = 6.048$, $p < .05$, indicating that the mean score for principals in the lower schools ($M = 1.55$, $SD = .165$) was significantly higher than principals in the upper schools ($M = 1.51$, $SD = .186$). However, the means are very close which indicates there may not be an educationally significant difference between the groups. The data presented in Table 28 indicate that principals in the lower performance group use more data tools in their efforts to increase mathematics scores on standardized assessments than principals in the higher performance group. The main effect of school size yielded an F ratio of $F(2, 146) = 2.339$, $p > .05$, indicating that the means for the school size groups did not differ significantly. The interaction effect between school performance and size was not significant.

A Look at the Impact of Years as Principal at Current School and Performance

A two-way ANOVA was conducted by the researcher to determine if a significant difference existed between the upper schools, lower schools, and years as principal at the school. Descriptive statistics for this analysis are presented in Table 29. A significant difference was found between the mean scores for years as principal at current school and rating of satisfaction data effectiveness in improving mathematics scores on standardized assessments. The main effect of years as principal yielded an F ratio of $F(2, 30) = 4.714$, $p < .05$, indicating that the

Table 27

Descriptive Statistics for Performance/School Size and Use of Data Tools

Performance	School Size	Mean	Std. Deviation	N
Lower	Rural	1.53	.172	49
	Suburban	1.58	.168	16
	Urban	1.61	.108	10
	Total	1.55	.165	75
Upper	Rural	1.47	.147	34
	Suburban	1.56	.192	35
	Urban	1.37	.390	3
	Total	1.51	.186	72
Total	Rural	1.50	.164	83
	Suburban	1.56	.184	51
	Urban	1.56	.213	13
	Total	1.53	.177	147

Table 28

Two-Way ANOVA for Performance/School Size and Use of Data Tools

Source	Type III Sum of Squares	*df*	Mean Square	*F*	Sig.
Corrected Model	.333(a)	5	.067	2.222	.055
Intercept	144.500	1	144.500	4822.382	.000
Performance	.181	1	.181	6.048	.015
School Size	.140	2	.070	2.339	.100
Performance* School Size	.092	2	.046	1.539	.218
Error	4.225	141	.030		
Total	347.926	147			
Corrected Total	4.558	146			

Note. a R Squared = .073 (Adjusted R Squared = .040).

Table 29

Descriptive Statistics for Performance/Years as Principal and Perceived Effectiveness of Satisfaction Data

Performance	Age	Mean	Std. Deviation	N
Lower	1-4 years	3.57	.63	7
	5-8 years	4.22	.51	3
	9+ years	4.17	.55	6
	Total	3.92	.63	16
Upper	1-4 years	3.08	1.13	4
	5-8 years	9.91	.53	7
	9+ years	4.00	.27	4
	Total	3.71	.75	15
Total	1-4 years	3.39	.83	11
	5-8 years	4.00	.52	10
	9+ years	4.10	.45	10
	Total	3.82	.69	31

means between principals with 1-4 years (M = 3.39, SD = .83), 5-8 years (*M* = 4.00, *SD* = .52), and 9+ years (*M* = 4.10, *SD* = .45) experience differed significantly. More specifically, the Tukey HSD revealed that the mean scores between principals with one-four years (*M* = 3.39, *SD* = .83) experience at their current school and principals with 9+ years (*M* = 4.10, *SD* = .45) differed significantly. The data presented in Table 30 suggest that principals with 9 or more years experience in the same school believe satisfaction data is more effective in improving standardized mathematics scores than principals with only one-four years experience in the same school. The main effect of performance yielded an *F* ratio of $F(2, 30) = 1.80$, $p > .05$, indicating that the mean score for lower schools (*M* = 3.91, *SD* = .63) did not differ significantly from that of upper schools (*M* = 3.71, *SD* = .755). The interaction effect was non-significant, $F(1, 30) = .856$, $p > .05$ (see Table 30).

Table 30

Two-Way ANOVA for Performance/Years as Principal and Effectiveness of Satisfaction Data

Source	Type III Sum of Squares	*df*	Mean Square	*F*	Sig.
Corrected Model	3.989(a)	5	.789	1.956	.121
Intercept	409.601	1	409.601	1004.210	.000
Performance	.735	1	.735	1.802	.191
Years as Principal	3.628	2	1.814	4.447	.022
Performance* Years	.128	2	.064	.157	.856
Error	10.197	25	.408		
Total	465.889	31			
Corrected Total	14.186	30			

Note. a R Squared = .281 (Adjusted R Squared = .137).

Analysis by Performance and Gender

A series of two factor ANOVAs were conducted to examine the main effects and interaction of gender and performance in each of the 11 sub-categories. Additionally, chi-square analysis was completed to determine whether significant differences existed between principal responses in the upper and lower schools, gender, and survey items in the sub-categories use of input data, outcomes data, process data, satisfaction data, and use of data tools. No significant differences were found between the upper and lower schools, gender, and use of data tools.

A two-way ANOVA was conducted by the researcher for performance, gender and the eleven sub-categories. One significant main effect was found for gender for perceived effectiveness of process data. Additionally, one significant interaction was found between performance and gender for the use of input data.

A significant difference was revealed between male and female principals for the perceived effectiveness rating of process data. The main effect of gender yielded an F ratio of $F(1, 42) = 6.895$, $p. < .05$ (see Tables 31 and 32), indicating that the mean score for females (M = 4.30, SD = .546) differed significantly from that of males (M = 3.85, SD = .561). It appears that female principals believe process type data to be more effective than male principals in increasing student achievement in mathematics on standardized assessments. Descriptive statistics are presented in Table 31.

Table 31

Descriptive Statistics for Principals' Effectiveness Rating of Process Data by Gender

Performance	Gender	Mean	Std. Deviation	N
Lower	Male	3.80	.50	17
	Female	4.67	.33	3
	Total	3.93	.56	20
Upper	Male	3.91	.63	15
	Female	4.17	.56	8
	Total	4.00	.61	23
Total	Male	3.85	.56	32
	Female	4.30	.54	11
	Total	3.97	.58	43

Table 32

Two-Way ANOVA for Performance/Gender and Effectiveness of Process Data

Source	Type III Sum of Squares	*df*	Mean Square	*F*	Sig.
Corrected Model	2.286	3	.762	2.453	.078
Intercept	469.060	1	469.060	1509.754	.000
Performance	.264	1	.264	.851	.362
Gender	2.142	1	2.142	6.895	.012
Performance* Gender	.631	1	.631	2.033	.162
Error	12.117	39	.310		
Total	691.778	43			
Corrected Total	14.403	42			

Note. $p<.05$.

Though no additional main effects for gender or performance were found, further two-way ANOVA indicated a significant interaction between gender and performance for the use of input data $F(1, 146) = 4.766$, $p < .05$ (see Tables 33 and 34). The nature of this interaction is shown in Figure 6. More specifically, it appears there is a significant difference in the mean cited use of input data between females ($M = 1.48$, $SD = .34$) in high and low ($M = 1.28$, $SD = .39$) performing schools. Though this difference is mathematically significant, the means likely to do not represent an educational significance. Descriptive statistics are presented in Table 33.

Table 33

Descriptive Statistics for Performance/Gender and Use of Input Data

Performance	Gender	Mean	Std. Deviation	N
Lower	Male	1.31	.31	58
	Female	1.28	.39	17
	Total	1.30	.33	75
Upper	Male	1.22	.30	59
	Female	1.48	.34	13
	Total	1.26	.32	72
Total	Male	1.26	.31	117
	Female	1.37	.38	30
	Total	1.28	.32	147

Table 34

Two-Way ANOVA for Performance/Gender and Use of Input Data

Source	Type III Sum of Squares	*df*	Mean Square	*F*	Sig.
Corrected Model	.803	3	.268	2.556	.058
Intercept	164.195	1	164.195	1567.358	.000
Performance	.073	1	.073	.697	.405
Gender	.333	1	.333	3.183	.077
Performance* Gender	.499	1	.499	4.766	.031
Error	14.981	143	.105		
Total	257.500	147			
Corrected Total	15.784	146			

Note. $p<.05$.

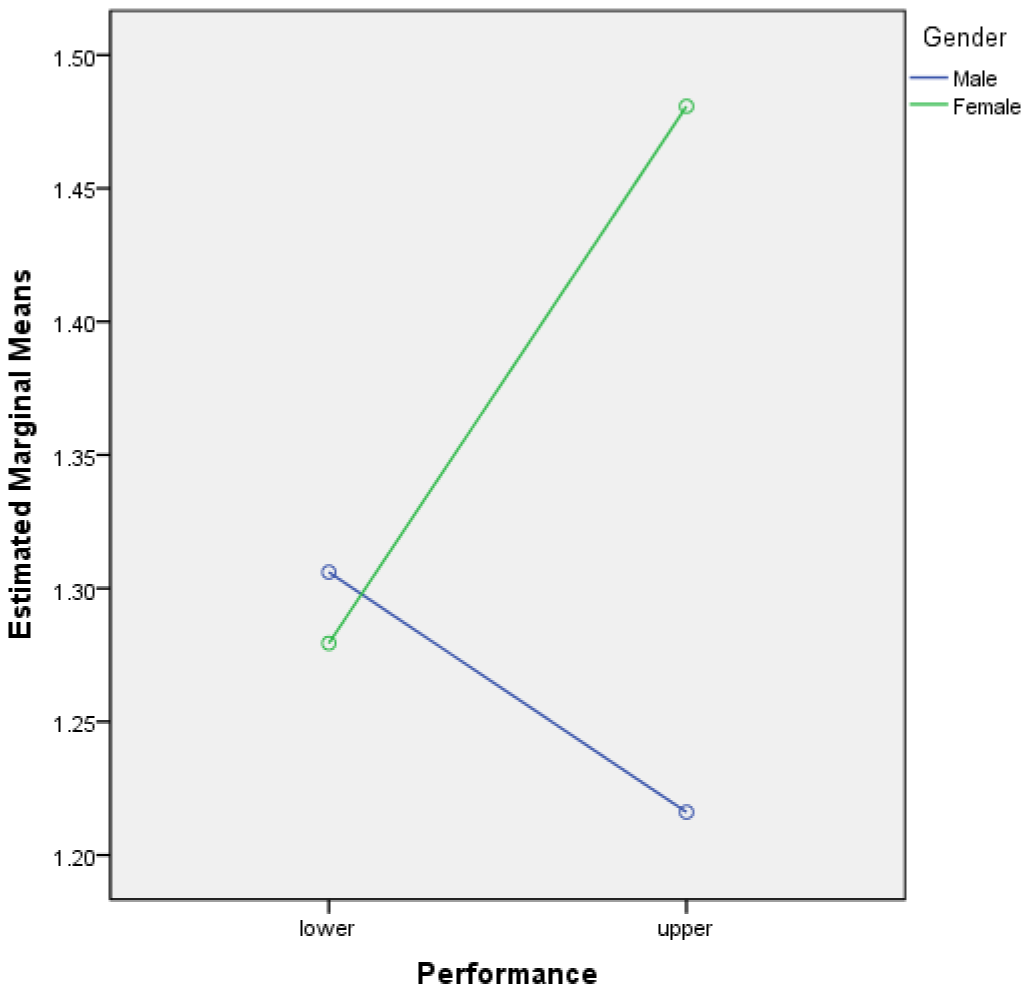

Figure 6. Plot of gender/performance and use of input data.

Chi-square revealed no significant differences were found between principal responses in the upper and lower schools, gender and 16 of 17 individual data items. A significant difference was revealed for one individual data item in the input data type sub-category. Table 35 shows a significant difference, χ^2 (1, N = 72) = 8.466, $p < .05$), between males and females in the upper performance group for the use of student language proficiency data. The results suggest that more female principals than males in the upper group use student language proficiency data to increase student achievement on standardized mathematics assessments.

Summary

Analysis shows that principals use a variety of data types and data tools though not in equal parts. Outcomes data is used most by principals in both performance groups. Principals in lower performing schools rate outcomes data as the most effective data type, while principals in the upper schools rate process data as more effective in their efforts to improve mathematics scores. While data tools designed specifically for use with outcomes type data had the highest cited usage among principals in both groups, principals in the upper performance group rate district and school created data tools as the most effective tools they use in efforts to increase mathematics scores. Principals in the upper and lower groups differ in the extent they use data systems and in their cited use of data tools. Additional differences exist between principals in high and low performing schools when looking at the impact of age and the years of experience principals have in their current school.

Table 35

Principals' Use of Student Language Proficiency Data

			No	Yes	Total	df	Sig.
Lower	Male	Count	47	11	58		
		Expected Count	45.6	12.4	58.0		
	Female	Count	12	5	17		
		Expected Count	13.4	3.6	17.0		
	Total	Count	59	16	75		
		Expected Count	59.0	16.0	75.0		
		Pearson Chi-Square Count				1	0.335
Upper	Male	Count	52	7	59		
		Expected Count	48.3	10.7	59.0		
	Female	Count	7	6	13		
		Expected Count	10.7	2.3	13.0		
	Total	Count	59	13	72		
		Expected Count	59.0	13.0	72.0		
		Pearson Chi-Square Count				1	0.004

Note. $p<.05$.

Data usage by principals in both the upper and lower group was similar in the sub-categories use of outcomes data, use of process data, use of satisfaction data, and use of input data. Slightly higher usage was cited by principals in the upper group for use of process type data. The analysis suggests that outcomes data is the type of data used by the most principals in both high and low performing schools followed closely by process type data, while fewer principals in both groups use satisfaction and input data types in their efforts to increase mathematics scores on standardized assessments. The individual data items of standardized test score data and data collected from local benchmark assessments in the outcomes type data sub-category had the highest identified use among principals in both groups followed by the individual item, data about curriculum needs in the process type data subcategory. Principals in both groups cite a greater use of outcomes type data than any other data type in their efforts to increase mathematics scores. Though outcomes type data is cited as used the most among principals in the upper group, principals in the upper group report a slightly higher mean effectiveness rating for process data in efforts to improve mathematics scores. Principals in the lower group rated outcomes type data as the most effective in their efforts to increase student achievement in mathematics on standardized assessments. Principals in both groups rely more on outcomes and process type data, and perceive these as more effective in their efforts to enhance student achievement on standardized assessments in mathematics.

Principals cited the highest usage of data tools that are designed specifically to manipulate and analyze outcomes type data. Though principals indicated the highest use for these types of tools, principals in the upper group reported a slightly higher mean rating for district and school created data tools. Principals in the lower group cited the 4Sight Benchmark Assessment tool as the most effective tool in their efforts to increase student achievement in mathematics. Though a high percentage of principals reported use of the Pennsylvania Value Added Assessment tool, the tool had the second lowest rating of effectiveness. The National Association for Educational Progress tool was rated as the least effective tool by principals in both groups.

Principals in the lower group had a slightly higher mean for their use of data systems. A series of one-way ANOVAs revealed no statistically significant differences between means in the 11 sub-categories for principals in the upper and lower performance groups. Chi-square analysis revealed statistically significant differences between means for principals in the upper and lower groups in their use of local benchmark assessment data and the use of the 4Sight Assessments/Success for All Foundation data tool. The results suggest that a greater number of principals in lower performing schools use local benchmark assessment data and the 4Sight Assessments/Success for All Foundation data tool. Chi-square analysis revealed a further difference between males and females in the upper group in the use of an individual data item. The analysis indicates that female principals use student language proficiency data more than male principals.

Significant differences were found in both the main effects of performance and age for the sub-category principals' use of data systems. Surprisingly, principals in the lower performance group use data systems to a greater extent than principals in the upper performance group. However, it must be noted that though the means are mathematically statistically difference, they are still very close and may not indicate a significant difference in educational practices. The main effect of age showed that principals 55+ years of age use data systems to a greater extent than principals 40-54 years of age. One additional two-way ANOVA revealed a significant difference in the main effect of performance for overall use of data tools. Principals in the lower performance group cited a higher use of data tools in their efforts to increase mathematics scores. Again, though these means are statistically different they are still very close and it should, therefore, be noted that they may not indicate significant difference in educational practices.

Additional findings revealed significant differences between age groups, years experience at current school, and school size. The findings indicate that principals 55+ years of age cited higher use of satisfaction data. Furthermore, principals in the low performing schools in this age group rate input and process type data as more effective in their efforts to increase student achievement in mathematics on standardized assessments. A look at the impact of the factors of principal years experience at their current school and performance revealed that principals with 9+ years of experience in their school perceive satisfaction data as more effective than principals with 1-4 years experience in their school.

Chi-square analysis by gender and performance showed significant differences between upper performance group males and upper performance group females in their use of student language proficiency. Significantly more females than males in the upper performance group cited use of student language proficiency data in efforts to increase mathematics scores on standardized assessments. Two-way ANOVA using the factors of performance and gender show that female principals rate the effectiveness of process data higher than male principals. Further two-way ANOVA showed that more female principals in the upper performance group reported use of input type data than females in the lower performance group.

In the following chapter, the researcher reviews the key findings of this study presented in this analysis. Additionally, conclusions and implications for practicing principals will be offered based on the findings of this study. Furthermore, areas for further study based on this research are identified.

CHAPTER 5

CONCLUSIONS

Introduction

The No Child Left Behind (NCLB) Act of 2001 has ushered in an era of high stakes testing designed to hold schools accountable for student performance in mathematics, reading, and writing. The legislation has provided real targets for student achievement with real consequences for failing to meet those targets. The requirements and expectations set forth by the law have significantly impacted teacher and principal perceptions of accountability as they feel the pressure to reach required levels of proficiency on standardized assessments (Davis, 2006; Johnson, 2006; LoGerfo, 2004). The pursuit of reaching established performance levels for proficiency has forced the development of a new skill set for school principals. This skill set includes the ability to collect and analyze data in order to support decision making by teachers and administrators that is designed to improve student performance. The ability to apply the skills of data use to increase student achievement has been recognized under the Interstate School Leaders Licensure Consortium (ISLLC) that drafted and adopted standards for school leaders. The ISLLC standards hold that all school leaders should have the knowledge of information sources, understanding of data collection, and proficiency in strategies for data analysis.

In the state of Pennsylvania, student proficiency levels in mathematics have lagged significantly behind proficiency levels in reading since the passage of NCLB (Pennsylvania Department of Education, 2009). Standardized assessment scores for Pennsylvania students mirror the scores of students on a national level. On average student performance on standardized mathematics assessments evidences a decline in mathematics proficiency as students matriculate from the elementary to the secondary level.

The effective use of data systems can lead to increased student achievement through multiple measures including standardized assessments (Bernhardt, 2007; Schmoker, 2006; Streiffer, 2004). Several studies have indicated that the use of data to support decision making can lead to school improvement (Alonzo, 2006; Armstrong & Anthes, 2001; Barry, 2006). Research also indicates that principals who are proficient in using data to inform their decisions have a significantly positive impact on student achievement (Englert, et al., 2004; Halverson, 2007; Marzano, 2005). On a national level this potential impact appears to be understood by principals as they have been steadily increasing their use of data to inform all types of decisions (Stecher & Hamilton, 2006; U.S. Department of Education, 2008). Though Pennsylvania principals have followed this trend by turning to the use of data systems to support decisions in their efforts to increase student achievement in mathematics, further research indicates that Pennsylvania principals are in the early stages in their development and use of data systems (Marsh, et al., 2006).

Researchers have repeatedly found that effective use of data, data tools, and data systems, are common components in high achieving schools (Armstrong & Anthes, Breiter & Light, 2006; Datnow, Park, & Wohlstetter, 2007). The purpose of this study was to analyze principals' use of data systems including various data types and tools to gain an understanding of what principals in higher performing schools do differently with data as compared to principals in lower performing schools. The study surveyed 147 secondary principals working in Pennsylvania public schools with highest percentage of students proficient and the lowest percentage of students proficient in grade 11 on the 2008-2009 PSSA mathematics assessment.

Principal responses were examined in two groups including the higher performance group ranked 1 through 150 and the lower ranked performance group ranked 151 through 300.

The study analyzed responses to 41 items in 11 different sub-categories including use of input data, use of outcomes data, use of process data, use of satisfaction data, perceived effectiveness of input data, perceived effectiveness of outcomes data, perceived effectiveness of process data, perceived effectiveness of satisfaction data, use of data systems, use of data tools, and perceived effectiveness of data tools. The instrument was designed by the researcher and included 11 items in the use of data systems sub-category that were used with permission from Englert (2004). Survey responses were examined to determine whether perceptions and use of data tools, data types, and data systems to enhance mathematics scores on standardized assessments differed in high and low performing schools.

The remainder of this chapter reviews the key findings of this research study. Limitations of this research are also addressed. The researcher presents conclusions based on the findings of this study including how these conclusions relate to the literature reviewed in previous chapters. Furthermore, implications for practicing principals based on this research are reviewed. Areas identified for further study are presented followed by a final summary.

Summary of Key Findings

This research focused on the perceptions related to the use of data types, data tools, and data systems between secondary principals in high and low performing schools as measured by the PSSA standardized mathematics assessment for 2008-2009. Perceptions between groups did not differ in the use of input data, use of process data, use of outcomes data, use of satisfaction data, perceived effectiveness of input data, perceived effectiveness of process data, perceived effectiveness of outcomes data, perceived effectiveness of satisfaction data, and perceived effectiveness of data tools used. Unexpected significant differences were found between performance groups in their use of data systems and overall use of data tools. Two-way ANOVA revealed that principals in the low performing schools perceived they use data systems to a greater extent and have a higher overall use of data tools than principals in high performing schools. Perhaps principals in the lower performing schools have worked to implement and increase their use of data driven decision making in their schools in response to poor student performance on standardized assessments. Some significant differences between principals in the high and low performing schools were revealed when individual survey items were analyzed using chi-square. In addition, two-way ANOVA revealed some significant differences between principals in the high and low performing schools when comparing age, years as principal at current school, and urban, suburban, and rural schools. Lastly, two factor chi-square analyses and further two-way ANOVA revealed some differences between gender and principals in high and low performing schools.

Principals' Use of Data Types and Individual Data Items

More Pennsylvania principals in high performing schools cited the use of outcomes type data (M = 1.63, SD = .027) over any other type as being used in efforts to increase mathematics scores, followed by process type data (M = 1.61, SD = 0.34), satisfaction type data (M = 1.46, SD = .035), and input type data (M = 1.26, SD = .033). However, principals in higher performing schools reported a higher mean use of only process type data than principals in lower performing schools. Furthermore, these principals rated process type data as the most effective type (M = 4.00, SD = 0.61) in guiding their efforts to increase mathematics scores followed by outcomes type data (M = 3.75, SD = 0.71), satisfaction type data (M = 3.71, SD = 0.75), and input type data (M = 3.21, SD = 1.09). Principals in the low performing schools reported the highest mean effectiveness rating for outcomes type data (*M* = 3.95, *SD* = 0.77), followed by process type data (*M* = 3.93, *SD* = 0.57), satisfaction type data (*M* = 3.92, *SD* = 0.63), and input data (*M* = 3.34, *SD* = 0.89).

More principals in both groups cited the use of standardized test score data and data collected from local benchmark assessments than any other individual data item in efforts to increase student achievement in mathematics on standardized assessments. In the lower group, 99% of principals cited use of data collected from local benchmark assessments compared to 76% of principals who reported use of this individual data item in the upper group. The percentage of principals who cited use of standardized test scores data to increase mathematics scores was much closer as 79% of principals in higher performing schools reported use, while 81% of principals in lower performing schools reported use. Data about curriculum needs followed these items closely in terms of use and effectiveness rating by both groups. Though a high percentage of principals in both groups reported the use of feedback data from teachers and teacher observation data, the effectiveness rating for both individual data items was comparatively low. Principals in both groups rated student language proficiency as the most effective individual data item in the input type data sub-category, while student gender data was rated as least effective.

Principals' Use of Data Tools

More principals in both groups cited the use of the Pennsylvania AYP data tool, the Emetric data tool, and the 4Sight data tool which are all designed to work with standardized test score data collected from state mandated assessments or gathered from local benchmark assessments. A greater number of principals in the lower performing schools cited use of these tools as 93% reported use of the Pennsylvania AYP data tool and 91% of principals in the lower group reported use of the 4Sight data tool. Though approximately half of the principals in both groups cited use of value added data, 79% of principals in the lower group and 69% of principals in the upper group reported use of the PVAAS data tool. The NAEP data tool was used by the lowest number of principals including 9% of principals in the lower performing schools who reported use and 13% of principals in the higher performing schools reported use.

Principals in the lower performing schools reported a higher mean effectiveness rating for the 4Sight data tool and rated the effectiveness of the tool higher than all other data tools. Though few principals in high and low performing schools cited use of district or school created data tools, these district and school created tools had the highest mean rating of effectiveness in the higher performing schools. Both groups cited a high use of the Emetric and Pennsylvania AYP data tools with over 80% of principals reporting use of each tool. However, the reported mean rating for effectiveness of these tools was lower than the mean rating for district developed tools, school developed tools, the 4Sight data tool, and Excel spreadsheets in both groups. The PVAAS data tool was given the same rating of effectiveness (M = 3.39, SD = 1.03) by both groups and ranked near the bottom ahead of only the NAEP data tool. The NAEP data tool was the only tool with a negative rating for effectiveness by one of the groups as rated by principals (M = 2.75, SD = .089) in the upper group.

Principals' Use of Data Systems

Principals were asked to rate the extent that they used data systems. Principals in the low performing schools reported the extent of their use with a mean of 3.94 (SD = 0.63) in their efforts to increase mathematics scores on standardized assessments. Principals in the higher performing schools reported a lower mean in the extent of their use of data systems (M = 3.78, SD = 0.58).

Principal responses in the high and low performing schools were analyzed for differences in 11 sub-categories using one-way ANOVAs. One-way ANOVA revealed no significant differences between the high and low performing schools and their use of outcomes type data, process type data, satisfaction type data, and input type data. However, follow-up analysis with chi-square on individual data items indicated a significant difference between principals in high and low performing schools in their reported use of data collected from local benchmark assessments, χ^2 (1, N = 147) = 16.967, p < .05). Principals in lower performing schools cited a higher use of data collected from local benchmark assessments.

Findings Using Two-Way ANOVA

Two-way ANOVA for size and performance showed a significant main effect for performance in the use of data tools, F(1, 146) = 6.048, p < .05. The mean use of data tools for principals in low performing schools (*M* = 1.55, *SD* = .16) was significantly higher than principals in the upper schools (*M* = 1.51, *SD* = .18). Though these means were statistically different, they are very close and may not indicate a true difference in educational practices between the groups. Principals in the lower performance group use more data tools in their efforts to increase mathematics scores on standardized assessments. Additional analysis of individual data tools using chi-square revealed a significant difference between principals in the high performing schools and principals in the low performing schools in their use of the 4Sight Assessments/Success For All Foundation data tool, χ^2 (1, N = 147) = 13.904, p < .05. Significantly more principals in the low performing schools use the 4Sight Assessments/Success For All Foundation data tool in their efforts to increase student achievement on standardized assessments in mathematics.

Principals in the lower performance group reported a higher mean use of data systems. Two-way ANOVA using the factors of performance and age for the sub-category use of data systems revealed that the main effect of performance differed significantly between the groups, F(1,146) = 4.112, $p < .05$. Surprisingly, principals in the low performing schools (M = 3.94, SD = 0.63) reported a perceived greater extent of use of data systems than principals in the higher performing schools (M = 3.78, SD = 0.58). Though mathematically statistically different, these means are very close and may not be an indication of significant differences in educational practices. The main effect of age showed further significant differences as principals 55+ years of age (M = 4.12, SD = .47) in high performing schools reported that they use data systems to a greater extent than principals 40-54 years old (M = 3.80, SD = .60). The main effect of age revealed an additional significant difference in the reported use of satisfaction data. The Tukey HSD follow up analysis indicated the average use of satisfaction data was significantly lower for principals 40-54 years old (M = 4.22, SD = 1.01) than principals 55+ years old (M = 4.85, SD = 1.02).

Additional two-way ANOVA for performance and age revealed significant interaction effects in two sub-categories including perceived effectiveness of input data and perceived effectiveness of process data. The Tukey HSD follow up analysis revealed that principals 55+ years (M = 4.31, SD = .23) old had a significantly higher mean perceived effectiveness rating of input data than principals 40-54 years old (M = 2.62, SD = .59) in low performing schools. A significant interaction effect was also revealed in the perceived effectiveness of process data in efforts to increase student achievement in mathematics on standardized assessments. The Tukey HSD follow up analysis showed that in the low performing schools, principals 55+ years old

(M = 4.88, SD = .19) perceived process type data to be more effective in increasing standardized mathematics scores than principals 40-54 year old (M = 3.70, SD = .45) and principals 25-39 years old (M = 3.81, SD = .41).

Two-way ANOVA for the factors of performance and years as principal at current school revealed a significant difference for the main effect of years as principal at current school in the mean effectiveness rating of satisfaction type data, $F(2, 30) = 4.714$, $p < .05$. Follow up analysis using the Tukey HSD showed that principals with 1-4 years (M = 3.39, SD = .83) experience at their current school and principals with 9+ years (M = 4.10, SD = .45) differed significantly. Principals with 9+ years experience at their current school perceived satisfaction type data to be more effective in efforts to increase student achievement in mathematics than principals with only 1-4 years experience at their current schools.

Some differences with respect to gender were found in the analysis. Two-way ANOVA for performance and gender showed a significant difference between female principals in low (M = 1.28, SD = .39) and high (M = 1.48, SD = .34) performing schools for their use of input data. More female principals in higher performing schools cited use of input type data in their efforts to improve student achievement in mathematics on standardized assessments than female principals in low performing schools. Further two-way ANOVA indicated that female principals (M = 4.30, SD = .546) believed process type data to be more effective in increasing mathematics scores than male principals (M = 3.85, SD = .561). Final chi-square analysis revealed a difference between male and female principals in the upper group for their cited use of student language proficiency data, χ^2 (1, N = 72) = 8.466, $p. < .05$). In the upper performance group, 46% of females indicated the use of student language proficiency data compared to 11% of males.

Limitations of the Study

The results of this study indicated that there are some significant differences between principals in low and higher performing schools in their use of data types, data tools, and data systems. However, the study does have some limitations including concern regarding self-reporting bias and confidence interval. At a 95% confidence level the overall response rate of 49% for this study has a confidence interval of 5.8% that is slightly higher than the targeted level of 5%. The response rate for the study was lower than the expected response of 56%. Additionally, the structure of this study was based on accurate self-reporting by the respondents. Therefore, the results may have been affected by self-reporting bias if principals exaggerated responses to portray themselves as better practitioners in the use of data to make decisions related to improving student achievement in mathematics.

There are additional items that should be noted as study limitations. The sample population was purposefully selected and excluded 54 of 602 secondary principals from two Pennsylvania school districts because their districts do not readily permit their participation in research studies. This exclusion may have adversely affected results of the study. The use of a monetary incentive to complete the survey may have introduced bias which adversely affected the results of the survey. Some research indicated that the use of monetary incentives introduces bias (Houston & Ford, 1976). However, more recent research has shown that offering incentives does not introduce bias and increases response rates (Singer, Groves, & Corning, 1999).

This research study was further limited by the scope of its examination. The instrument used in this research study sought to find if principals used specific data items in several subgroups. Additionally, it was designed to find principals' perceived effectiveness of data types and items. The instrument was not designed to understand specifically how the data items and types were used in practice by principals in their efforts to increase student achievement in mathematics. This may be a limiting factor in fully understanding the responses analyzed in this study.

Conclusions

The results of this research showed that a variety of data, data tools, and data systems are being used in varying degrees by principals to make decisions that guide their efforts to improve student achievement in mathematics on standardized assessments like the PSSA. The main focus of this research study was to understand what principals in higher performing schools do differently than principals in low performing schools in their use of data driven decision making to increase student achievement on standardized assessments in mathematics. Some findings revealed that data types, data systems, and data tools are being used differently by principals in higher performing schools.

Principals in both groups indicated the use of different types of data in varying degrees and high use of data systems. Principals in both groups also reported that they are using the data tools made available to them by the Pennsylvania Department of Education. These results were similar to research finding of Hamilton, et al. (2007) that showed principals and mathematics teachers reported increased use of data to direct school improvement in mathematics. Hamilton, et al. (2007) attributed the increased data use to new accountability standards defined under NCLB.

The results of this research study paralleled the findings of previous studies that showed principals in both high and low performing schools use outcomes type data sources more than other data types to inform their decision making including that standardized assessment data from state exams and benchmark assessments are used most often by principals to guide decision making (Englert, 2004; Stecher & Hamilton, 2006). Though a difference was not found between groups for their use of data types in this study, the findings of Datnow, Park, and Wohlstetter (2007) indicated that high performing schools use a variety of data sources in decision making to guide instruction in subjects like mathematics. Principals in both groups indicated use of a variety of data sources. Furthermore, principals in both groups perceived outcomes, process, and satisfaction data types to be most effective in their efforts to increase mathematics scores, while input type data is perceived as less effective.

Based upon the results of this research study, it can be concluded that principals in both groups are using the data tools provided to them by the Pennsylvania Department of Education. These findings reflected the research of McCaffrey and Hamilton (2007) that also indicated Pennsylvania principals are using the data tools provided to them. However, the findings of this study indicated that principals who have school or district created data tools in the higher performance groups rated those tools to be more effective than state provided data tools in their efforts to increase mathematics scores. Principals in low performing schools use the 4Sight/Success for All Foundation data tool significantly more and rate it as the most effective tool in their efforts.

Principals in high and low performing schools report high use of data systems. Unexpectedly, the results of this research indicated that principals in the higher performing schools use data systems to a lesser extent than principals in the lower performing schools. Additionally, principals in the upper performance group cited less overall use of data tools than principals in the lower group. In similarity to the principals in the low performing schools, principals in the higher performing schools cited high use of outcomes type data. Though principals in the upper performance group rated the two most used individual data items of locally collected benchmark assessment data and standardized test scores data as effective or more effective than principals in the lower performing schools, principals in the upper performance group rated process data to be more effective than outcomes data. Marsh, et al. (2006) noted that case study research indicated that educators in high performing schools systematically examine data on school and classroom practices. Teachers in these schools indicated that they found classroom assessments closely aligned to daily lesson objectives to be more effective and timely in guiding their instruction than data collected from progress or benchmark type assessments. Upper performance group principals in this research study rated outcomes type data similarly to satisfaction data in terms of effectiveness. Principals in the upper performance group may see a greater value in process data (e.g., data about curriculum needs, data best teaching practices, data about program effectiveness) than outcomes type data in their approach to improving student achievement in mathematics.

Less use of data tools and data systems by principals in higher performing schools may reflect that they perceive less pressure to increase test scores. Furthermore, the results suggest that high use of data tools, benchmark assessment data, and data systems by the principal may only be one component to improving student achievement in mathematics. Perhaps, the expectation to use data in efforts to improve student achievement in mathematics may be shared by principals with teachers to a greater degree in the higher performing schools. Student achievement may be attributable to systems developed by these principals that are producing desired results. Principals in the upper performance group rated data tools that were school or district created as the most effective in their efforts to increase mathematics scores. These school and district developed tools may be more effective in producing results for their specific schools.

As discussed previously, research has shown that NCLB has influenced accountability in schools. The results of this study reflected the well documented pressure principals feel to increase student achievement on standardized assessments (Davis, 2006; Johnson, 2006). Given the high stakes attached to testing under the federal mandate, it is not surprising principals in both high and low performing schools cited the highest use of outcomes type data. Principals in lower performing schools are likely feeling greater pressure to increase scores under NCLB, which was reflected in their significantly higher use of data tools, locally collected benchmark assessment data, and data systems. Findings show these principals viewed the use of outcomes type data (e.g., locally collected benchmark assessment data) and associated tools (e.g., 4Sight/Success for All Foundation data tool) over other types of data as more effective in efforts to increase student achievement in mathematics on tests like the PSSA. More specifically, it is clear that principals in the lower performing schools believe that locally collected benchmark data is the most effective individual data item that they use to guide their efforts to increase student scores in mathematics. Furthermore, principals in the low schools cited significantly higher usage of the 4Sight/Success for All Foundation data tool than principals in the higher schools. They also rated it as the most effective data tool suggesting that they believe the use of this tool will lead to improved student achievement in mathematics on standardized assessments. Research has indicated that the use of these types of data tools can lead to improved standardized test scores (Breiter & Light, 2006; Light, et al., 2004). However, research has also cautioned against the over use of a single data source (Koretz, 2003).

Principals in the lower schools may feel a greater need to keep a close and frequent watch on student progress toward achieving mathematics standards through use of locally collected benchmark assessment data so that necessary interventions may be implemented prior to taking state level standardized assessments (e.g., tutoring students in benchmark identified deficiencies) at the end of a school year. Greater use of data tools and data systems by principals in the lower schools suggests that though their focus is on the use of benchmark data and the 4Sight data tool, they are exploring many options in their efforts to increase their student scores in mathematics on standardized assessments.

Further findings in this study showed that principals 55+ years old rate input data as more effective than principals 40-54 years old. Older principals may believe characteristics like race and socioeconomic status are more useful than their younger counterparts because they are employed in schools that are held accountable for results by these sub-categories under NCLB. Principals 55+ years old also cited a greater usage of satisfaction data than their counterparts in the 40-54 year old age bracket, which implies their maturity level has led them to value the opinions of others to a greater degree in their efforts to increase student performance in mathematics on standardized assessments. Further findings showed that principals with 9+ years experience at their current school believed satisfaction data was more effective in efforts to increase mathematics scores than principals with only 1-4 years experience. As principals gain greater experience and understanding over time in one school, they may rely and trust opinions of others to a greater degree.

Recommendations for Practitioners

Though principals in low performing schools clearly view the use of data systems as a solution to increasing student achievement, their overuse of a single data source and data tool may be misguided. Use and effectiveness ratings by principals in the upper performance group reflect a more balanced approach to increasing student achievement in mathematics. High usage of a single data source is in contrast to good practice for data driven decision making and systems thinking. Over reliance on specific data source that is gathered at one point in time (e.g., PSSA assessments) to make decisions is bad practice for data driven decision making (Bernhardt, 2007; Marsh, et al., 2006; Streiffer, 2004). Data collected from local benchmark assessments and state standardized tests have limited content coverage. Systems thinking is the foundation for data driven decision making and it requires consideration of many variables across several levels (Senge, 2000). Not using other types of data at appropriate levels makes real causes difficult to pinpoint in systems. Using multiple sources of data provides a more balanced approach to decision making and reduces the possibility that any one individual data item will become corrupted in the system resulting in poor decision making that ignores relevant additional factors (Koretz, 2003).

Though significant differences were not found between the groups in the use and rating of process type data between the performance groups, principals in the upper performance group rated this type of data as most effective in their efforts to increase mathematics scores on standardized assessments. These principals clearly believed benchmark data and an associated benchmark data tool (e.g., the 4Sight/Success for All Foundation data tool) to be less effective than principals in low performing for increasing student achievement in mathematics. They may have worked to triangulate findings to provide a more balanced approach to data driven decision making as noted previously. Furthermore, the study findings indicated that principals in the upper groups may have developed systems using data to produce desired results standardized assessments in mathematics. This may be why principals in the upper performance groups report less use of data systems and data tools.

Principals should approach the development of a data system in their school starting with fundamental questions to understand what data will be used, and how it will be used. These fundamental questions as posed by Bernhardt (1998) should be a starting point for schools establishing a culture of continuous improvement. Bernhardt contended that the remaining six questions are influenced by the answer to the first.

1. What do you expect students to know and be able to do by the time they leave the school? (Standards)
2. What do you expect students to know and be able to do by the end of each year? (Benchmarks)
3. How well will students be able to do what they want to do with the knowledge and skills they acquire by the time they leave school? (Performance)
4. Do you know why you are getting the results you are getting? Do you know why are you not getting the results you want?
5. What would your school and educational processes look like if your school were achieving its purpose, goals, and expectations for student learning?
6. How do you want to use the data you gather? (p. 9)

As noted by Bernhardt (2007), it is through this process that a true systems approach can be developed for the use of data in schools because it will necessarily involve examination of problems through different lenses that are comprised of different types of data.

Recommendations for Further Study

As discussed earlier, the results of this study indicated secondary principals in Pennsylvania rely heavily on outcomes type data to make decisions designed to improve student standardized test scores in mathematics on assessments like the PSSA. Further results showed that principals in lower performing schools viewed the use of outcomes type data, data tools, and data systems as key to increasing student achievement on standardized assessments in mathematics. Based on the results of this study, the following are recommendations for further study that may provide more insight into the use of data to increase student achievement in mathematics as measured by standardized assessments:

1. Qualitative or quantitative studies should be conducted in other states in high and low performing schools to determine whether these results are representative of national perceptions of principals in their efforts to increase mathematics scores.
2. Qualitative or quantitative studies should be conducted with elementary and middle school principals to exam whether these results are similar to the perceptions of principals across grade levels in their efforts to increase student achievement in mathematics.
3. Qualitative or quantitative studies should be conducted to examine secondary mathematics teachers' perceptions of how they use data types, tools, and data systems to guide instruction.
4. Qualitative or quantitative studies should be conducted to determine the utility of various data types across levels (e.g., types most useful for a given purpose for teachers, for building administrators, for district administrators) in efforts to increase standardized scores in mathematics.
5. Qualitative or quantitative studies should be conducted to examine the quality of data driven decisions being made in schools by teachers and principals in their efforts to increase mathematics scores.
6. Qualitative or quantitative studies should be conducted to examine the effects of increased use of outcomes type data to guide decision making directed at increasing student standardized mathematics assessment scores.

7. Qualitative or quantitative studies should be conducted to examine how specific data items and data types are used in the decision making processes of principals across their areas of job responsibility.

Final Summary

Previous research has found that the No Child Left Behind Act has influenced principals' perceptions of accountability (Davis, 2006). Requirements under the act that mandate accountability through standardized assessments have forced principals to develop skills using data. Clearly, virtually all principals are using data to inform decision making related to standardized assessments in some measure. Results from this study show that principals in both high and low performing schools use various types of data, data tools, and data systems. The results of this study suggest that principals in low performing schools see the use of data as the cure for poor student achievement in mathematics as they feel the pressure of high stakes accountability under the No Child Left Behind Act. Pressure to increase student achievement may have resulted in greater use of data tools and systems as principals in low performing schools search for solutions to poor student achievement. Though the No Child Left Behind Act has fostered the development of data skills by principals, unintended consequences like overuse of single data sources may also be a result of this influence.

It may be that principals in higher performing schools have found a more balanced approach in their efforts to increase student achievement in mathematics on standardized assessments. As suggested by Bernhardt (2007), these principals may have developed or had a pre-existing school culture that supports the use of data to guide instruction, which they have adjusted over time in the continuous process of improvement. These adjustments may have resulted in increased efficiency of the data system used in their schools that uses only essential data and tools, while discarding superfluous data.

It is clear that principals see the use of data systems as a means to improving student achievement. The results of this study indicate that principals in high and low performing schools over rely on standardized assessment data to make decisions. This is considered back practice for data driven decision making. However, if student achievement continues to be measured by a single source like state standardized assessments, principals' use of data in struggling schools may continue to become significantly distorted from what is considered good practice. These principals appear to be resigned to a narrow focus by zeroing in on student weaknesses identified using benchmark data instead of taking a systems thinking approach oriented toward long term enduring improvement.

References

Abelmann, C., Elmore, R., Even, J., Kenyon, S., & Marshall, J. (1999). *When accountability knocks, will anyone answer?* Philadelphia, PA: CPRE Publications.

Alonzo, T. L. (2006). *Closing achievement gaps through data-driven decision making in a success for all school* (Doctoral dissertation). Available from Proquest database. (UMI No. 3237780)

Amrein-Beardsley, A. (2008). Methodological concerns about the education value-added assessment system. *Educational Researcher, 37*, 65-75.

Anderson, C. M. (2005). *Evaluation of the effects of longitudinal tracking of student achievement to assess school quality*. Los Angeles: University of Southern California. (UMI No. 3180487)

Armstrong, J., & Anthes, K. (2001). *Identifying key factors, conditions, and policies that support schools' use of data in decision making and school improvement: Summary of findings*. Retrieved November 15, 2007 from Education Commission of the States website http://www.ecs.org/clearinghouse/30/69/3069.htm

Babbie, E. (2004). *The practice of social research*. Belmont, CA: Wadsworth.

Bali, V., & Alavarez, M. (2003). Schools and educational outcomes: What causes the race. *Social Science Quarterly, 84*, 485-507.

Barry, M. J. (2006). *A school's use of data for teaching and learning: A case study of data impact on instruction in an urban school* (Doctoral dissertation). Available from Proquest database. (UMI No. 3237164)

Bernhardt, V. L. (1998). *Data analysis for comprehensive school improvement.* Larchmont, NY: Eye on Education.

Bernhardt, V. L. (2000). Intersections: New routes open when one type of data. *Journal of Staff Development, 21*, 33-37.

Bernhardt, V. L. (2007). *Translating data into information to improve teaching and learning*. Larchmont, NY: Eye on Education.

Bigger, S. L. (2006). *Data-driven decision making within a professional learning community: Assessing the predictive qualities of curriculum-based measurements to a high-stakes, state test of reading achievement at the elementary level* (Doctoral dissertation). Available from Proquest database. (UMI No. 3209996)

Buchanan, B. (2004). Defining adequate yearly progress. *American School Board Journal, 191*, 10-13.

Breiter, A., & Light, D. (2006). Data for school improvement: Factors for designing effective information systems to support decision-making in schools. *Educational Technology & Society*, 9(3), 206-217.

Carey, K. (2004). The real value of teachers: Using new information about teacher effectiveness to close the achievement gap. *Thinking K-16, 8*(1), 1-42.

Cole, N. (2001). The new faces of fairness. *Journal of Educational Measurement, 38,* 369-373.

Creswell, J. (1998). *Qualitative inquiry and research design: Choosing among the five*. Oakmont, CA: Sage.

Data Quality Campaign. (2009). *Creating a longitudinal data system to improve student achievement.* Retrieved November 15, 2007, from Data Quality Campaign website www.dataqualitycampaign.org

Datnow, A., Park, V., & Wohlstetter, P. (2007). *Achieving with data: How high performing school systems use data to improve instruction for elementary students*. Retrieved from the University of Southern California website http://www.usc.edu/dept/education/cegov/focus/education_reform/publications/other

Davis, C. (2006). *Secondary school principals' perceptions of the value and impact of high-stakes tests on content and mode of instruction* (Doctoral dissertation). Available from Proquest database. (UMI No. 2352630)

Dembosky, J., Pane, J., Barney, H., & Christina, R. (2005). *Data-driven decision making in southwestern Pennsylvania school districts* (Working paper). Santa Monica, CA: RAND Education. Retrieved September 27, 2009, from http://www.rand.org/pubs/working_papers/2006/RAND_WR326.pdf

Dougherty, C. (2002). *A policymaker's guide to the value of longitudinal student data*. Retrieved September 30, 2004, from Education Commission of the States: http://www.ecs.org/clearinghouse/40/21/4021.htm

Duke, D. (2004). *The turnaround principal: high stakes leadership*. Retrieved from National Association of Elementary School Principals (NAESP): http://www.naesp.org/resources/2/Principal/2004/S-Op12.pdf

Education Commission of the States. (2002). *Data-driven decision making: No child left behind brief*. Retrieved September 30, 2004, from the National School Boards Association website http://www.nsba.org/MainMenu/Governance/DataDrivenDecisionMaking.aspx

Englert, K., Fries, D., Goodwin, B., Martin-Glenn, M., & Michael, S. (2004). *Understanding how principals use data in a new environment of accountability*. Aurora, CO: Mid-continent Research for Education and Learning.

Farkas, S., Johnson, J., Duffett, A., Foleno, T., & Foley, P. (2001). *Trying to stay ahead of the game: Superintendents and principals talk about school leadership*. Retrieved February 15, 2010, from http://www.publicagenda.com/research/pdfs/ahead_of_the_game.pdf

Foley, E., Mishook, J., Thompson J., Kubiak, M., Supovitz, J., Rhude-Faust, M. K. (2008). *Beyond test scores: Leading indicators for education.* Retrieved November 27, 2009 from the Annenberg institute: http://www.annenberginstitute.org/index.php

Fullan, M. (2007). *The new meaning of educational change*. New York, NY: Teachers College Press.

Fuhrman, S. H. (1999). *The new accountability CPRE policy brief*. Philadelphia, PA: Consortium on Policy Research in Education. (ERIC Document Reproduction Service No.ED429336)

Gelman, A. (2004). *Analysis of variance---why it is more important than ever*. Retrieved November 1, 2010 from Columbia University: http://www.stat.columbia.edu/~gelman/research/published/banova7.pdf

Glod, M. (2009). *Nation's report card sees gains in elementary and middle*. Retrieved February 1, 2010 from Washington Post: http://www.washingtonpost.com/wp-dyn/content/article/2009/04/28/AR2009042801244.html

Gross, M. (2007). *Preparing future leaders: Principals' perceptions of the interstate school leaders licensure consortium standards and knowledge indicators* (Doctoral dissertation). Available from Proquest database. (UMI No. 3315396)

Halverson, R., Grigg J., Prichett R., & Thomas, C. (2005, July). *The new instructional leadership: Creating data-driven instructional systems in schools*. Paper prepared for the annual meeting of the National Council of Professors of Educational Administration, Washington, DC.

Halverson, R., Prichett, R., & Watson, J. (2007, May). *Formative feedback systems and the new instructional leadership* (WCER Working Paper No. 2007-3). Retrieved February 7, 2010 from University of Wisconsin-Madison, Wisconsin Center for Education Research: http://www.wcer.wisc.edu/publications/workingPapers/papers.php

Hamilton, L., Stecher, B., Marsh J., McCombs, J., Robyn, A., Russell, J., … Barney, H. (2007). *Standards based accountability under No Child Left Behind: Experiences of teachers and administrators in three states*. Santa Monica, CA: RAND Education.

Hamilton, L., Berends, M., & Stecher, B. (2005, April). *Teachers' responses to standards-based accountability*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Hamilton, L. (2004). *Assessment as a policy tool*. Santa Monica, CA: RAND Education.

Hammond, D. (1997). *Toward a science of synthesis: the heritage of general systems theory* (Doctoral dissertation). Available from Proquest database. (UMI No.9803216)

Hawley, W. D. (1988). Missing pieces of the educational reform agenda: Or, why the first, and second waves may miss the boat. *Education Administration Quarterly*, *4*, 416-437.

Heck, R. H. (2000). Examining the impact of school quality on school outcomes and improvement: A value-added approach. *Education Administration Quarterly, 36*, 51-73.

Heritage, M., & Chen, E. (2005). Why data skills matter in school improvement. *Phi Delta Kappan International, 86*, 707-710.

Hoff, D. (2001). *U.S. Ranks among world's best and worst readers*. Retrieved January 27, 2010 from Reading Rockets website http://www.readingrockets.org/article/328

Hopkins, L. (2007). *PISA results released: U. S. students fall behind in international education study*. Retrieved January 28, 2010, from Suite101.com website http://international-education.suite101.com/article.cfm/us_students_left_behind

Houston, M., & Ford, N. (1976). *Broadening the scope of methodological research on mail surveys*. Retrieved October 10, 2010, from www.jstor.org website http://www.jstor.org/pss/3151023

Huck, S. (2003). *Reading Statistics and Research*. Boston, MA: Allyn & Bacon.

Izumi, L. (2009). *How to improve national mathematics scores*. Retrieved January 30, 2010, from New York Times website http://roomfordebate.blogs.nytimes.com/2009/10/15/how-to-improve-national-mathematics-scores/

Jennings, J. (1999). *Why national standards and tests?* Thousand Oaks, CA: Sage.

Johnson, D. D. (2006). *A study of the effects of test-driven accountability on perceptions of principals in the Dallas independent school district* (Doctoral dissertation). Available from Proquest database. (UMI No. 3238588)

Kelly, S., & Monczunski, L. (2007). Overcoming the volatility in school-level gain scores: A new approach to identifying value added with cross-sectional data. *Educational Researcher, 3,* 279-287.

Killon, J., & Bellamy, G. T. (2000). On the job: Data analysts focus school improvement efforts. *Journal of Staff Development, 21*, 27-32.

Kim, J., & Sunderman, G. L. (2004). *Large mandates and limited resources: State response to the No Child Left Behind Act and implications for accountability*. Cambridge, MA: The Civil Rights Project at Harvard University.

Klein, D. (2003). *A brief history of American K-12 mathematics education in the 20th century*. Retrieved January 28, 2010, from the California State University at Northridge website http://www.csun.edu/~vcmth00m/AHistory.html

Koretz, D. (2003). Using multiple measures to address perverse incentives and score inflation. *Educational Measurement: Issues and Practice, 22*(2), 18-26.

Leithwood, K., Louis, K., Anderson, S., & Wahlstrom, K. (2004). *How leadership influences student learning*. New York, NY: Wallace Foundation.

Light, D., Wexler, D., & Heinze, J. (2004, April). *How practitioners interpret and link data to instruction: Research findings on New York City schools' implementation of the Grow Network*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Littlejohn, S. W. (1983). *Theories of human communication*. Belmont, CA: Wadsworth.

LoGerfo, L. F. (2004). *Does responsibility matter and who is responsible?: Exploring the influence and development of teacher responsibility for childrens' welfare* (Doctoral dissertation, The University of Michigan, Ann Arbor). (UMI No. 3137880)

Mandinach, E., Honey, M., & Light, D. (2006, April). *A theoretical framework for data-driven decision making*. Paper presented at the annual meeting of AERA, San Francisco.

Mandinach, E., Rivas, L., Light, D., Heinze, C., & Honey, M. (2006). *The impact of data-driven decision making tools on educational practice: A systems analysis of six school* *districts*. Retrieved December 2, 2008 from the Center for Children and Technology website http://cct.edc.org/admin/publications/speeches/Data_AERA06.pdf

Mann, D., & Shakeshaft, C. (2003). In God we trust; all others bring data. *School Business Affairs, 1*, 19-21.

Marsh, J., Pane, J., & Hamilton, L. (2006). Making sense of data-driven decision making in education. Santa Monica, CA: RAND Education.

Marzano, R. (2003). *What works in schools: Translating research into action*. Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R., Waters, T., & McNulty, B. (2005). *School leadership that works*. Alexandria, VA: Association for Supervision and Curriculum Development.

McCaffrey, D. F., & Hamilton, L. S. (2007). *Value added assessment in practice: Lessons from the Pennsylvania valued added assessment system pilot project*. Santa Monica, CA: The Rand Corporation.

McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003). *Evaluating value-added models for teacher accountability*. Santa Monica, CA: The Rand Corporation.

Meadows, H. (2008). *Principals' use of data in small and rural school districts of Florida* (Doctoral dissertation). Retrieved from http://etd.fcla.edu/WF/WFE0000101/Meadows_Hubert_O'Neal_200805_EdD.pdf

Means, B., Gallagher, L., & Padilla, C. (2007). *Teachers' use of student data systems to improve instruction.* Washington, DC: U.S. Dept. of Education, Office of Planning, Evaluation and Policy Development.

Means, B., Padilla, C., DeBarger, A., & Bakia, M. (2009). *Implementing data-informed decision making in schools: Teacher access, supports, and use*. Washington, DC: Department of Education, Office of Planning, Evaluation, and Policy Development.

Merlino, F., Bernotsky, L., Feldman, J., & Rui, N. (2009). *The financial and non-financial factors associated with 11th grade student academic proficiency in mathematics and reading in Pennsylvania.* Retrieved February 24, 2010 from Mathematics Science Partnership of Greater Philadelphia (MSPGP): http://www.mspgp.org/library/0_25536500%201245358115_725688498.pdf

Meyer, R. H. (1997). Value-added indicators of school performance: A primer. *Economics of Education Review, 16*(3), 283-301.

Mizes, S., Fleece, E., & Roos, C. (1984). Incentives for increasing return rates: Magnitude levels, response bias, and format. *Public Opinion Quarterly, 48*(4),794-800.

Murphy, J. (2002). How the ISLLC standards are reshaping the principalship. *Principal, 82*(1), 22-26.

National Center for Education Statistics. (2002). *Building an automated student record system: A step-by-step guide for local and state education agencies*. Retrieved September 30, 2004, from National Center for Education Statistics website http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2000324

Nyankori, R. J. (2005). *Does certification of elementary school teachers matter? The effects of certification status on instructional practices and on the mathematics and reading achievement of first grade public school students*. College Park: University of Maryland.

Newmann, F. M., King, M. B., & Rigdon, M. (1997). Accountability and school performance: Implications from restructuring schools. *Harvard Educational Review, 67*(1), 41-74.

No Child Left Behind Act of 2001. (2001). Retrieved September 3, 2004, from U. S. Department of Education http://www.ed.gov/nclb/landing.jhtml?src=mr

North Carolina Department of Public Instruction (NCDPI)/Accountability Services. (2006). *North Carolina's proposal to pilot the use of a growth model for AYP purposes.* Retrieved from the North Carolina Department of Public Instruction website http://www.dpi.state.nc.us/docs/accountability/reporting/abc/2005-06/ncaypgrowthproposal041606.pdf

O'Day, J., Clune, B., Mason, S., & Thiel, C. (1999). *Looking for school improvement in Milwaukee: An example of embedded research*. Retrieved October 8, 2003, from Wisconsin Center for Education Research website: http://www.wcer.wisc.edu/mps/AERAoday.htm

Pennsylvania Department of Education. (2008). *Mathematics and reading technical report*. Retrieved September 12, 2009, from Pennsylvania Department of Education website: http://www.able.state.pa.us/a_and_t/lib/a_and_t/2008_Mathematics_and_Reading_Technical_Report.pdf

Pennsylvania Department of Education. (2007). *Pennsylvania information management system*. Retrieved September 12, 2009, from Pennsylvania Department of Education website: http://www.edportal.ed.state.pa.us/portal/server.pt/community/pims-pasecureid/1839

Pennsylvania Department of Education. (2008). *Data tools in a standards aligned system.* Retrieved December 14, 2008, from Pennsylvania Department of Education website: http://www.pde.state.pa.us/a_and_t/lib/a_and_t/DataTools9-15-08.pdf

Pennsylvania Department of Education. (2008). *Navigating and interpreting PVAAS reports.* Retrieved December 14, 2008, from Pennsylvania Department of Education website: http://www.pde.state.pa.us/a_and_t/lib/a_and_t/User_Manual.pdf

Pennsylvania Department of Education (2009). *PSSA and AYP results.* Retrieved January 28, 2010, from Pennsylvania Department of Education website: http://www.education.state.pa.us/portal/server.pt/community/school_assessments/7442

Phillips, L. M., Norris, S. P., Osmond, W. C., & Maynard, A. M. (2002). Relative reading achievement: A longitudinal study of 187 children from first through sixth grades. *Journal of Educational Psychology, 94*, 3-13.

Popham, W. J. (1987). The merits of measurement-driven instruction. *Phi Delta Kappan*, 679-682.

Popham, W. J. (2001). *The truth about testing.* Alexandria, VA: ASCD.

Pulliam, J., & Van, Patton, J. (1999). *History of education in America*. Upper Saddle River, NJ: Prentice Hall.

Ramanathan, A. K. (2007). *Wave of the future or passing fancy? The expanded federal role in the oversight and enforcement of the No Child Left Behind Act and the Individuals with Disabilities Education Act* (Doctoral dissertation). Available from Proquest database. (UMI No. 3271699)

Rescorla, L., & Rosenthal, A. (2004). Growth in standardized ability and achievement test scores from 3rd to 10th grade. *Journal of Educational Psychology, 96*, 85-96.

Rubin, D. B., Stuart, E. A., & Zanutto, E. L. (2004). A potential outcomes view of value-added assessment in education. *Journal of Educational and Behaviorial Statistics, 29*(1), 103-116.

Sanders, W. L. (2000). Value-added assessment from student achievement data: Opportunities and hurdles. *Journal of Personnel Evaluation in Education, 14*(4), 329-339.

Sanders, W. L., & Horn, S. P. (1994). The Tennessee value-added assessment system (TVAAS): Mixed-model methodology in educational assessment. *Journal of Personnel Evaluation in Education, 8*, 299-311.

Sanders, W. L., & Horn, S. P. (1998). Research findings from the Tennessee value-added assessment system (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education, 12*(3), 247-256.

Schmidt, W. (1998, June). *Changing mathematics in the U.S.: Policy Implications from the Third International Mathematics and Science Study*. The Annual Conference of the National Council of Teachers of Mathematics, Washington, DC.

Senge, P. (1990). *The fifth discipline.* New York, NY: Doubleday.

Senge, P., Cambron-Mcabe, N., Lucas, T., Smith, B., Dutton, J., Kleiner, A. (2000). *Schools that learn.* New York, NY: Doubleday.

Singer, E., Groves, R., & Corning, A. (1999). Differential incentives: Beliefs about practices, perceptions of equity, and effects on survey participation. *Public Opinion Quarterly, 63*, 251-260.

Stecher, B., & Hamilton, L. (2006). *Using test-score data in the classroom* (Working paper). Santa Monica, CA: RAND Education. Retrieved October 22, 2007, from Pennsylvania Department of Education website http://www.rand.org/pubs/working_papers/2006/RAND_WR375.pdf

Stein, M. (2003). *Making sense of the data: Overview of the k-12 data management and systems market.* Boston, MA: Eduventures.

Stone, C., & Lane, S. (2001). Consequences of a state accountability program: Examining relationships between school performance gains and teacher styles, and school variables. *Applied Measurement in Education, 16*, 1-26.

Streifer, P. (2004). *Tools and techniques for effective data-driven decision making*. Toronto, Canada: Scarecrow Education.

Teddlie, C., & Reynolds, D. (2000). *The international handbook of school effectiveness research*. London, England: Falmer.

Tennessee Department of Education. (2006). *Proposal to the U. S. Department of Education: NCLB growth model pilot program*. Retrieved October 22, 2007, from Tennessee Department of Education website http://www.tennessee.gov/education/nclb/doc/NCLB%20GrowthModelProposal.pdf

Torrence, V. (2002). *Principals use of data: a national perspective* (Doctoral dissertation). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.10.6809&rep=rep1&type=pdf

U. S. Department of Education, Office of Planning, Evaluation and Policy Development. (2008). *Teachers use student data systems to improve instruction: 2005 to 2007*. Washington, DC: Author.

Vandervoort, L. G., Amrein-Beardsley, A., & Berliner, D. C. (2004). National board certified teachers and their students' achievement. *Education Policy Analysis Archives, 12(46).* Retrieved from the Education Policy Analysis Archives website http://epaa.asu.edu/epaa/v12n46/v12n46.pdf

Wayman, J., Stringfield, S., & Yakimowski, M. (2004). *Software enabling school improvement through analysis of student data* (CREPSPAR Technical Report No. 67). Baltimore, MD: Center for Research on the Education of Students Placed at Risk (CRESPAR).

Willingham, W. (2002). Grades and test scores: Accounting of observed differences. *Journal of Educational Measurement, 39,* 1-38.

Wilson, M. (2004). *Towards coherence between classroom assessment and accountability*. Chicago, IL: University of Chicago Press.

Appendices

APPENDIX A

Cover letter Informed Consent Form

May 8, 2010

Dear Colleague,

As a secondary school principal in a public school I know that PSSA assessments are a significant concern for you, particularly at this time of the year. **The pressure on you as a principal to demonstrate high performance on PSSA assessments is very real, particularly in math where the percentage of Pennsylvania students scoring proficient or above on the PSSA in grade 11 lags significantly behind the averaged scores of students in grades 3 through 8 over each year the test has been given.** Principals at all levels have been strongly encouraged by the PA Dept. of Education to utilize data systems to improve student performance. Your input regarding your use of data systems is valuable information that could help other principals in their efforts to meet the demands of the No Child Left Behind Act.

My name is Josh Williams, I am a secondary public school principal currently working in Pennsylvania. I am also a doctoral student at Indiana University of PA working under the supervision of Dr. Sue Rieg (srieg@iup.edu, 724-357-2416.) I am inviting you to participate in a research project study on how secondary principals use data systems to improve student achievement in mathematics as measured by the PSSA. Listed below is a web address to a questionnaire that asks a variety of questions about how principals perceive the process of using data systems to improve math scores on standardized tests. I hope you will take the time to complete this questionnaire. It should not take more than 10 minutes of your time.

Indiana University of Pennsylvania supports the practice of protection of human subjects participating in research. This project has been approved by the Indiana University of Pennsylvania Institutional Review Board for the Protection of Human Subjects (Phone: (724) 357-2223). There are no known risks or discomforts associated with this research. Please be aware that even if you agree to participate in this survey study, you are free to withdraw at any time by emailing me at jwilliams@smasd.org and you may do so without penalty. Although your participation is solicited, it is strictly voluntary.

Your name will never be placed on a survey and your name will not in any way be associated with any of the findings. All information obtained will be kept confidential and incorporated into group data. Please type the web address into your Internet browser to complete the survey by **May 15, 2010.** Your submission of a completed survey implies consent as a participant in this study.

Survey web address: www.XXXXXXX

If you have any questions or require additional information, please do not hesitate to contact me at (814) 781-2123 or jwilliams@smasd.org

I appreciate your time and cooperation and look forward to receiving your completed survey.

Sincerely,

Josh Williams
Doctoral Candidate
Indiana University of PA
Administration and Leadership Studies

Sue A. Rieg, Ed. D.
Professor
Indiana University of PA

APPENDIX B

Follow-up Email

May 30, 2010

Approximately two weeks ago you should have received a letter of invitation to complete a survey seeking your input regarding your use of data systems to improve student achievement on standardized tests for mathematics. This survey was sent to fellow principals currently practicing in Pennsylvania to gain their insights on their use of data driven decision making to increase student achievement in mathematics as measured by standardized assessments.

If you have already completed and returned the survey, thank you. If not, please do so today. Although your participation is solicited, it is strictly voluntary, as indicated in the previous mailing. Your completion of the survey will be greatly appreciated!

Survey web address: www.XXXXXXX

Thanks again for taking the time to complete the survey so it can be included in my doctoral research at Indiana University of Pennsylvania. If you have any questions or require additional information, please do not hesitate to contact me at (814) 781-2123 or jwilliams@smasd.org

Sincerely,

Josh Williams
Doctoral Candidate
Indiana University of PA
Administration and Leadership Studies
Stouffer Hall
Indiana, PA 15705

Sue A. Rieg, Ed. D.
Professor
Indiana University of PA
724-357-2416

APPENDIX C

Follow-up Email (21 day)

June 8, 2010

Dear Colleague,

Approximately four weeks ago, you should have received a letter of invitation to complete survey at the web address below seeking your input regarding your use of data systems to improve student achievement on standardized tests for mathematics. As of today, I have not received your completed survey. I would very much appreciate your feedback in the results.

This purpose of the survey was to gain insights of principals' use of data driven decision making to increase student achievement in mathematics as measured by standardized assessments. Your input is critical. Although your participation is solicited, it is strictly voluntary. Please consider returning the survey as soon as possible.

Survey web address: www.XXXXXXX

If you have any questions or require additional information, please do not hesitate to contact me at (814) 781-2123 or jwilliams@smasd.org

As indicated in previous mailings, thank you very much for your assistance in this research project!

Sincerely,

Josh Williams
Doctoral Candidate
Indiana University of PA
Administration and Leadership Studies
Stouffer Hall
Indiana, PA 15705

Sue A. Rieg, Ed. D.
Professor
Indiana University of PA
724-357-2416

APPENDIX D

Josh Williams

From: Michele Bloom [bloom@eyeoneducation.com] **Sent:** Mon 4/20/2009 2:43 PM
To: Josh Williams
Cc:
Subject: RE: permission
Attachments:

Permission is granted at no charge. Please make sure to include a full bibliographic reference.

Thank you for your interest in our pubications.

Michele Bloom

Assistant to the President

Eye On Education

6 Depot Way West

Larchmont, NY 10538

(888) 299-5350

From: Josh Williams [mailto:jwilliams@smasd.org]
Sent: Monday, April 20, 2009 2:46 PM
To: editor@eyeoneducation.com
Subject: FW: permission

I am uncertain of the process I must follow for this but here goes:

I am seeking permission to properly cite and use a scanned image of figure 3.1 on p. 25 from

Translating Data into Information to Improve Teaching and Learning, by Victoria L. Bernhardt, 2007, Larchmont, NY Eye on Education. Copyright 2007 Eye on Education, Inc.

APPENDIX E

Josh Williams

From:	Daniel Light [dlight@edc.org]	**Sent:**	Tue 4/21/2009 1:50 PM
To:	Josh Williams		
Cc:			
Subject:	Re: dddm conceptual framework		
Attachments:			

Dear Josh,

Thank you for your kind words, I will share them with my colleagues.

I would love to ask what you like about the framework in particular? I get a lot of requests from actual educators to cite our work and I have always wondered about the attraction. I don't mean to ask you to write a long answer, so please ignore if you don't have a short response.

Daniel

On 4/21/09 1:49 PM, "jwilliams@smasd.org" <jwilliams@smasd.org> wrote:

Dr. Light,

Thank you very much. I would like to add that I've appreciated your work in relation to mine as a principal as we continue to develop our own processes for utilizing student data.

Josh W.

From: Daniel Light [mailto:dlight@edc.org]
Sent: Tuesday, April 21, 2009 11:14 AM
To: Josh Williams
Subject: Re: dddm conceptual framework

Hi Josh,

Please feel free to cite our work and use the graphic in your dissertation. I can't remember the specific image attached so I am not sure where it came from. There is a slightly different image in the Ch 2., so I would suggest using that citation, even if you use this jpeg.

The version presented at AERA is different and is a much simpler image.

Daniel Light, Ph.D.

On 4/21/09 10:46 AM, "jwilliams@smasd.org" <jwilliams@smasd.org> wrote:
Dr. Light,

APPENDIX F

You replied on 4/21/2009 3:43 PM.

Josh Williams

From: Lauren Skrabala [skrabala@rand.org] **Sent:** Tue 4/21/2009 2:42 PM
To: Josh Williams
Cc:
Subject: Re:
Attachments:

Josh,
Permission is granted to use the figure from RAND document OP-170 in your dissertation, per your email below. Please indicate that the figure is reprinted with permission.

Please let me know if you have any questions.

Many thanks,
Lauren

Lauren Skrabala
Permissions
RAND Corporation

From: Josh Williams <jwilliams@smasd.org>
Date: Mon, 20 Apr 2009 15:26:42 -0500
To: <skrabala@rand.org>

I am sending this to request permission to use the graphic below contained in the cited paper. I would like to use it as part of my literature review on data driven decision making for my dissertation.

Josh Williams

Graduate Student at Indiana University of PA

Principal

Saint Marys Area High School, St. Marys, PA

APPENDIX G

MCREL

4601 DTC Blvd., Suite 500 • Denver, CO 80237
303.337.0990 • Fax: 303.337.3005 • www.mcrel.org

February 18, 2010

Josh Williams
977 South St. Marys Road
St. Marys, PA 15857

Permission is hereby granted to Josh Williams to use, adapt and cite in his doctoral dissertation for Indiana University of Pennsylvania the following material which is copyrighted by McREL:

- Assessment and Accountability Principal Survey

We request a standard scholarly citation to this material along with the statement "Used by permission of McREL."

We understand your dissertation will not be commercially published. This permission is limited to the material and purpose stated. Prior written permission is required for any additional uses.

Sincerely,

Linda Brannan

Linda Brannan
Lead Consultant

Appendix H

Principals' Use of Data Systems to Enhance Student Achievement in Mathematics Survey

The use of data by principals to inform decisions has become increasingly important since the passage of the No Child Left Behind Act. This survey is designed to assess the extent that you engage in practices using data to improve student achievement in math at your school. It should take about 10 minutes to complete. You may navigate forward and back through the survey using the navigation buttons at the bottom left.

Participation is voluntary. There are no consequences should you choose not to participate. All required actions have been taken to ensure respondent anonymity.

1. Please indicate your gender.

Male	Female
O	O

2. Please indicate your age range.

- O 25-39 years
- O 40-54 years
- O 55+ years

3. How many years have you worked as a principal in your current school?

- O 1-4 years
- O 5-8 years
- O 9+ years

4. Please indicate your school district's size:

- O Rural (0-1,999 students)
- O Suburban (2000-6,999 students)
- O Urban (7,000+ students)

5. Listed below are several types of data. Please click once on each individual type of data that you use to enhance student mathematics scores on the PSSA assessment. (After clicking on them once, selections will turn green, you may click a second time to de-select any data type.)

- DATA ABOUT THE EFFECTIVENESS OF SCHOOL PROGRAMS (e.g. tutoring programs)
- STUDENT LANGUAGE PROFICIENCY DATA
- FEEDBACK DATA FROM TEACHERS(e.g. their opinions on how programs designed to enhance student performance are working)
- TEACHER OBSERVATION DATA
- REPORT CARD DATA
- STANDARDIZED TEST SCORE DATA
- BENCHMARK ASSESSMENT DATA (e.g. 4Sight Assessments)
- STUDENT GENDER DATA
- FEEDBACK DATA FROM STUDENTS (e.g. opinions on whether a program or class helped them perform better)
- STUDENT SOCIOECONOMIC DATA
- VALUE ADDED ASSESSMENT DATA
- STUDENT ETHNICITY DATA
- DATA ABOUT BEST PRACTICES FOR INSTRUCTION
- TEACHER GENERATED AUTHENTIC ASSESSMENT DATA
- DATA ON PERCEPTIONS OF THE LEARNING ENVIRONMENT (e.g. student, teacher, parent opinion data regarding student preparedness)
- STUDENT ATTENDANCE DATA
- DATA ABOUT CURRICULUM NEEDS (e.g. examining how curriculum matches state standards and eligible content)

6. Please rate the extent that **teacher observation data** has been effective in enhancing student achievement in mathematics in your school as measured by standardized assessments.

NOT EFFECTIVE | ○ ○ ○ ○ ○ | VERY EFFECTIVE

7. Please rate the extent that **student ethnicity data** has been effective in enhancing student achievement in mathematics in your school as measured by standardized assessments.

NOT EFFECTIVE | ○ ○ ○ ○ ○ | VERY EFFECTIVE

8. Please rate the extent that **student language proficiency data** has been effective in enhancing student achievement in mathematics in your school as measured by standardized assessments.

NOT EFFECTIVE | ○ ○ ○ ○ ○ | VERY EFFECTIVE

9. Please rate the extent that **feedback data from teachers** has been effective in enhancing student achievement in mathematics in your school as measured by standardized assessments.

NOT EFFECTIVE | ○ ○ ○ ○ ○ | VERY EFFECTIVE

10. Please rate the extent that **student attendance data** has been effective in enhancing student achievement in mathematics in your school as measured by standardized assessments.

NOT EFFECTIVE | ○ ○ ○ ○ ○ | VERY EFFECTIVE

11. Please rate the extent that **data about best practices for instruction** has been effective in enhancing student achievement in mathematics in your school as measured by standardized assessments.

NOT EFFECTIVE | ○ ○ ○ ○ ○ | VERY EFFECTIVE

12. Please rate the extent that **data about curriculum needs** has been effective in enhancing student achievement in mathematics in your school as measured by standardized assessments.

NOT EFFECTIVE | O O O O O | VERY EFFECTIVE

13. Please rate the extent that **student gender data** has been effective in enhancing student achievement in mathematics in your school as measured by standardized assessments.

NOT EFFECTIVE | O O O O O | VERY EFFECTIVE

14. Please rate the extent that **standardized test score data** has been effective in enhancing student achievement in mathematics in your school as measured by standardized assessments.

I

NOT EFFECTIVE | O O O O O | VERY EFFECTIVE

15. Please rate the extent that **value added assessment data** has been effective in enhancing student achievement in mathematics in your school as measured by standardized assessments.

NOT EFFECTIVE | O O O O O | VERY EFFECTIVE

16. Please rate the extent that **data about school program effectiveness** has been effective in enhancing student achievement in mathematics in your school as measured by standardized assessments.

NOT EFFECTIVE | O O O O O | VERY EFFECTIVE

17. Please rate the extent that **student feedback data** has been effective in enhancing student achievement in mathematics in your school as measured by standardized assessments.

NOT EFFECTIVE | O O O O O | VERY EFFECTIVE

18. Please rate the extent that **data collected from benchmark assessments (e.g. 4Sight Assessments)** has been effective in enhancing student achievement in mathematics in your school as measured by standardized assessments.

NOT EFFECTIVE | O O O O O | VERY EFFECTIVE

19. Please rate the extent that **data on perceptions of the learning environment** has been effective in enhancing student achievement in mathematics in your school as measured by standardized assessments.

NOT EFFECTIVE | O O O O O | VERY EFFECTIVE

20. Please rate the extent that **teacher generated authentic assessment data** has been effective in enhancing student achievement in mathematics in your school as measured by standardized assessments.

NOT EFFECTIVE | O O O O O | VERY EFFECTIVE

21. Please rate the extent that **report card data** has been effective in enhancing student achievement in mathematics in your school as measured by standardized assessments.

NOT EFFECTIVE | O O O O O | VERY EFFECTIVE

22. Please rate the extent that **student socioeconomic data** has been effective in enhancing student achievement in mathematics in your school as measured by standardized assessments.

NOT EFFECTIVE | O O O O O | VERY EFFECTIVE

To what extent do you use data to inform your work as a principal in the following areas?

23. to realign instruction so that essential tested math curriculum is taught before students are assessed.

To no extent | ○ ○ ○ ○ ○ | To a great extent

24. to establish common outcome goals for student achievement in math amongst school staff.

To no extent | ○ ○ ○ ○ ○ | To a great extent

25. to identify school instructional strengths and weaknesses in math.

To no extent | ○ ○ ○ ○ ○ | To a great extent

To what extent do you use data to inform your work as a principal in the following areas?

26. to align school mathematics curriculum to state standards.

To no extent | ○ ○ ○ ○ ○ | To a great extent

27. to help develop school improvement plans designed to enhance student achievement in mathematics.

To no extent | ○ ○ ○ ○ ○ | To a great extent

28. to identify teacher instructional strengths and weaknesses in efforts to enhance student achievement in mathematics in your school.

To no extent | ○ ○ ○ ○ ○ | To a great extent

29. To what extent do you as a principal use state assessment data to monitor the progress of your school in mathematics?

To no extent | ○ ○ ○ ○ ○ | To a great extent

30. To what extent do you as a principal evaluate teachers on their ability to use data to inform their classroom practices in mathematics?

To no extent | ○ ○ ○ ○ ○ | To a great extent

31 To what extent do you believe teachers in your school meet the expectations for the use of data to improve student achievement in mathematics?

To no extent | ○ ○ ○ ○ ○ | To a great extent

32. To what extent does your school use data to track the performance of individual students over time in mathematics?

To no extent | ○ ○ ○ ○ ○ | To a great extent

33. To what extent does your district require you to meet data driven performance goals (e.g. reduce dropout rates by X%, increase student achievement in mathematics by X%)?

To no extent | ○ ○ ○ ○ ○ | To a great extent

34. To what extent do you as a principal ensure that data are used to differentiate instruction in mathematics in your school?

To no extent | ○ ○ ○ ○ ○ | To a great extent

35. To what extent do you as a principal use multiple sources of data to evaluate student progress in mathematics during the school year?

To no extent | ○ ○ ○ ○ ○ | To a great extent

36. Listed below are several tools for using data. Please click once on each individual tool that you use to enhance student mathematics scores on the PSSA assessment. (After clicking on them once, selections will turn green, you may click a second time to de-select any data type.)

- Pennsylvania Value Added Assessment System (PVAAS): PVAAS is a statistical analysis system that uses longitudinal data of students' performances on the PSSA assessments. (https://pvaas.sas.com)
- 4Sight Assessments/Success For All Foundation Assessing to Learn: The PA 4Sight Benchmarks are aligned to the PSSA and provide an estimate of student performance on the PSSA as well as diagnostic subskill data to inform classroom instruction decisions and professional development efforts. (http://members.successforall.net)
- eMetric: Data Interaction for Pennsylvania Student Assessments by eMetric are designed to provide quick, easy,and secure access to student performance results on the Pennsylvania System of School Assessment (PSSA). (https://pssa.emetric.net)
- NAEP: The National Assessment of Educational Progress, also known as "The Nation's Report Card," is the only nationally representative and continuing assessment of what America's students know and can do in various subject areas. (http://http://nces.ed.gov/nations reportcard)
- PA AYP: District and school reports that contain Adequate Yearly Progress (AYP) targets and results for the most recent year, the last two years, and next year's targets; and discussion of results for parents and educators. (http://paayp.emetric.net)
- Software Based Student Management System (i.e. electronic grade book, SASIxp, PowerSchool)
- Excel Spreadsheets
- A System Devised By Our District
- A System Devised By Our School

37. Please rate the extent that **Pennsylvania Value Added Assessment System (PVAAS)** data tool has been effective in enhancing mathematics scores.

NOT EFFECTIVE | ○ ○ ○ ○ ○ | VERY EFFECTIVE

38. Please rate the extent that **eMetric** data tool has been effective in enhancing mathematics scores.

NOT EFFECTIVE | ○ ○ ○ ○ ○ | VERY EFFECTIVE

39. Please rate the extent that **4Sight Assessments/Success For All Foundation** data tool has been effective in enhancing mathematics scores.

NOT EFFECTIVE | ○ ○ ○ ○ ○ | VERY EFFECTIVE

40. Please rate the extent that **PA AYP** data tool has been effective in enhancing mathematics scores.

NOT EFFECTIVE | ○ ○ ○ ○ ○ | VERY EFFECTIVE

41. Please rate the extent that **The National Assessment of Educational Progress (NAEP)** data tool has been effective in enhancing mathematics scores.

NOT EFFECTIVE | ○ ○ ○ ○ ○ | VERY EFFECTIVE

42. Please rate the extent that the **System Devised By Your School** data tool has been effective in enhancing mathematics scores.

NOT EFFECTIVE | ○ ○ ○ ○ ○ | VERY EFFECTIVE

43. Please rate the extent that the **System Devised By Your District** data tool has been effective in enhancing mathematics scores.

NOT EFFECTIVE | ○ ○ ○ ○ ○ | VERY EFFECTIVE

44. Please rate the extent that the **Excel Spreadsheets** data tool has been effective in helping you use data to enhance student achievement in mathematics.

NOT EFFECTIVE | ○ ○ ○ ○ ○ | VERY EFFECTIVE

45. Please rate the extent that the **Software Based Student Management System (i.e. electronic grade book, SASIxp, PowerSchool)** data tool has been effective in enhancing mathematics scores.

NOT EFFECTIVE | ○ ○ ○ ○ ○ | VERY EFFECTIVE