The Effects of the Toyota Production System on Student Academic Performance

Danny Ray Webb

Follow this and additional works at: https://dsc.duq.edu/etd

Recommended Citation
THE EFFECTS OF THE TOYOTA PRODUCTION SYSTEM ON STUDENT ACADEMIC PERFORMANCE

A Dissertation
Submitted to Duquesne University
School of Education

Duquesne University

In partial fulfillment of the requirements for
the degree of Doctor of Education

By
Danny Ray Webb

May, 2008
Copyright by

Danny Ray Webb

2008
DUQUESNE UNIVERSITY
SCHOOL OF EDUCATION
INTERDISCIPLINARY DOCTORAL PROGRAM FOR EDUCATIONAL LEADERS

Dissertation

Submitted in Partial Fulfillment of the Requirements For the Degree of Doctor of Education (Ed.D.)

Presented by:

Danny R. Webb
B.S., Industrial Arts/Technology Education, California University, 1987
M.Ed., Secondary Principal, Shippensburg University, 1995

January 17, 2008

THE EFFECTS OF THE TOYOTA PRODUCTION SYSTEM ON STUDENT ACADEMIC PERFORMANCE

Approval

_____________________________________________, Chair
James E. Henderson, Ed.D.
Professor of Educational Leadership and
Director, Interdisciplinary Doctoral Program for Educational Leaders
Duquesne University School of Education

_____________________________________________, Member
Helen C. Sobehart, D.A.
Associate Provost/Academic Vice President
Duquesne University

_____________________________________________, Member
Carol Parke, Ph.D.
Associate Professor, Department of Foundations & Leadership
Duquesne University

Program Director
James E. Henderson, Ed.D.
Professor of Educational Leadership and
Director, Interdisciplinary Doctoral Program for Educational Leaders
Duquesne University School of Education
THE EFFECTS OF THE TOYOTA PRODUCTION SYSTEM ON STUDENT ACADEMIC PERFORMANCE

By

Danny Ray Webb

May, 2008

Dissertation Supervised by Professor, James E. Henderson, EdD.

Abstract

Can schools use relevant data to enhance student test scores and teacher efficacy by using the Toyota Production System model? Components of this study evaluated fifth-grade student math Pennsylvania System of School Assessment scores, 4Sight Benchmark Assessment scores, and changes in teacher efficacy, as indicated by the Ohio State Teacher Efficacy Scale. The independent variables used to stimulate change are the Toyota Production System design for instruction and teacher training using Toyota Production System principles. Results did not indicate significant statistical changes in the Toyota Production System group’s Pennsylvania Systems of School Assessment scores, or 4Sight scores, as compared to the control group. Teacher efficacy for the Toyota Production System group did show trends of improvement when compared to control group; however, the results were not
statistically significant. This study creates an excellent foundation from which to base future studies of Toyota Production System in education; as the results from the descriptive portion and teacher interview indicate considerable amounts of learning for the Bedford School District and the implementation teacher.
ACKNOWLEDGMENTS

The author is grateful to those persons whose assistance and support have helped to make this dissertation a reality.

- Sincere thanks are extended to the members of the dissertation committee, Dr. James E. Henderson, Chairperson, Dr. Helen C. Sobehart, Dr. Carol S. Parke, and Mrs. Christina Dixon. Their expertise and counsel have truly made this a better product.

- Special thanks are due to Dr. Glenn E. Thompson for support and guidance as a mentor and friend.

- Special appreciation is expressed Dr. Patrick E. Crawford and the Bedford School Board for their support throughout this program.

- Particular thanks are also due my colleagues in the Interdisciplinary Doctoral Program for Educational Leaders who have helped me to perform at the highest level possible.

- Thanks are offered to those members of the Bedford Area School District who participated in this study. Without their help and hard work, this research study would not have been possible.

- Special thanks goes out to my family and friends who have sacrificed significant amounts of time throughout this process.
ACKNOWLEDGMENTS (cont.)

- Deserving more gratitude than he knows is my son, Chase, who has displayed patience beyond his years throughout this dissertation process and who has always helped me to keep that process in proper perspective.

- My deepest gratitude is reserved for my wife, Michelle. Without her patience, support, understanding, and love this project would not have been possible. I am grateful for the addition of our second son, Haydan John Webb. Thank you.
# TABLE OF CONTENTS

Chapter 1 ..............................................................................................................1  
Introduction ........................................................................................................1  
Research Questions ..............................................................................................5  
Research Hypothesis ............................................................................................6  
Identification of Variables ....................................................................................6  
Significance of Problem .......................................................................................7  
Operational Definitions .......................................................................................11  
Assumptions .........................................................................................................13  
Limitations ...........................................................................................................14  
Chapter 2 ............................................................................................................15  
Review of Selected Related Literature .................................................................15  
Necessity of Data ..................................................................................................16  
  Concerns Regarding Data Use ...........................................................................20  
  Application of Data ...........................................................................................23  
Other Factors Affecting Student Success ..............................................................25  
Toyota Production System ..................................................................................29  
TPS in Other Sectors .............................................................................................35  
TPS in Education ................................................................................................43  
PSSA ....................................................................................................................49  
4Sight Benchmark Assessments .........................................................................52  
Ohio State Teacher Efficacy Scale ....................................................................55
TABLE OF CONTENTS (cont.)

Summary........................................................................................................59

Chapter 3.......................................................................................................62

Introductory Statement..................................................................................62

Target Population..........................................................................................64

Method of Sampling......................................................................................64

Stimulus Materials.........................................................................................65

Measurement Devices....................................................................................69

PSSA..............................................................................................................69

4Sight Predictive Benchmark Assessments..................................................71

Ohio State Teacher efficacy Scale..............................................................72

Data Collection Methods............................................................................74

Research Design and Procedures...............................................................75

TPS Implementation Process.......................................................................78

Chapter 4.......................................................................................................83

Results..........................................................................................................83

Comparison of TPS to Control Group on Demographic Variables.............83

Student Performance....................................................................................86

PSSA Results...............................................................................................87

4Sight Benchmark Assessment Results......................................................89

Teacher Efficacy Results..............................................................................92
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classroom Teacher Interview</td>
<td>98</td>
</tr>
<tr>
<td>Research Findings</td>
<td>103</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>106</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>106</td>
</tr>
<tr>
<td>Discussion</td>
<td>108</td>
</tr>
<tr>
<td>TPS Implementation Discussion</td>
<td>109</td>
</tr>
<tr>
<td>Conclusions</td>
<td>113</td>
</tr>
<tr>
<td>Implications</td>
<td>114</td>
</tr>
<tr>
<td>Recommendations</td>
<td>119</td>
</tr>
<tr>
<td>References</td>
<td>122</td>
</tr>
<tr>
<td>Appendices</td>
<td>133</td>
</tr>
<tr>
<td>Appendix A: Summer Work Days</td>
<td>134</td>
</tr>
<tr>
<td>Appendix B: Model Line Plan</td>
<td>136</td>
</tr>
<tr>
<td>Appendix C: Classroom Model Design</td>
<td>138</td>
</tr>
<tr>
<td>Appendix D: Help Line Flowchart</td>
<td>140</td>
</tr>
<tr>
<td>Appendix E: Problem Identification Process</td>
<td>142</td>
</tr>
<tr>
<td>Appendix F: Problem Log</td>
<td>144</td>
</tr>
<tr>
<td>Appendix G: OSTES</td>
<td>146</td>
</tr>
<tr>
<td>Appendix H: PSSA/4Sight Comparison</td>
<td>149</td>
</tr>
<tr>
<td>Appendix I: Interview Questions</td>
<td>151</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Descriptive Statistics for PSSA Analysis</td>
<td>88</td>
</tr>
<tr>
<td>2.</td>
<td>Summary Results for PSSA Analysis</td>
<td>88</td>
</tr>
<tr>
<td>3.</td>
<td>Descriptive Statistics for 4Sight Analysis</td>
<td>90</td>
</tr>
<tr>
<td>4.</td>
<td>ANOVA Source of Variance Table for 4Sight Analysis</td>
<td>91</td>
</tr>
<tr>
<td>5.</td>
<td>Post Hoc Test Results for 4Sight Analysis</td>
<td>91</td>
</tr>
<tr>
<td>6.</td>
<td>Descriptive Statistics for Teacher Efficacy Analysis</td>
<td>93</td>
</tr>
<tr>
<td>7.</td>
<td>OSTES Group Mean Question Analysis</td>
<td>93</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. PSSA Math Mean Comparison</td>
<td>89</td>
</tr>
<tr>
<td>2. 4Sight Math Mean Comparison</td>
<td>91</td>
</tr>
</tbody>
</table>
CHAPTER ONE
INTRODUCTION

The No Child Left Behind Act (NCLB) is the latest national effort to bring “more accountability,” a favorite catchphrase of policymakers, to public education (Renter & Hamilton, 2003). Differing from past federal legislation, this law aims to hold public schools accountable for student performance. State-determined achievement tests increasingly serve as the centerpieces of state accountability systems (Popham, 2003). Under NCLB, test scores are used to determine whether schools are making “adequate yearly progress” in raising student achievement. Schools that do no make adequate yearly progress, according to their state’s achievement benchmarks, will be identified as “needing improvement,” even if they do not receive federal funds (Renter & Hamilton, 2003).

Just as medical tests help diagnose and treat patients, rigorous and meaningful education assessments can help ensure the academic growth of students (Gandal & McGiffert, 2003). School leaders and teachers face the challenge of utilizing data derived from local testing, state-mandated testing, and other credible sources to improve student achievement. In any education lexicon these days, the term data is inarguably one of the most positively loaded nouns (Popham, 2003). Data, upon which informed decisions can be made, offers school leaders a reliable source.

The most appropriate leader “is one who can lead others to lead themselves” (Manz & Sims Jr., 1999, p. 213). This new concept of leadership is referred to as “Super Leadership.” Focusing on what is truly important becomes essential for
organizations to be successful. Collins (2001) refers to the “hedgehog” concept, where leaders can see what is essential and ignore the rest. Applying the hedgehog concept in an educational organization would require school leaders to focus on what is essential. The hedgehog concept works when an organization and its leaders share the answers to the following three questions: At what can you be the best in the world? What drives your economic engine? About what are you deeply passionate (Collins, 2001)? Unlike businesses that produce physical items and expect to have the best product with consistent results, schools are challenged to educate students who possess varied ability levels that will never become consistent or standardized. What is the single concept, or hedgehog, that a school system can utilize to drive its organization from good to great?

Few employees are capable of executing highly effective self-leadership from the moment they enter a job situation (Manz & Sims Jr., 1999). According to Manz and Sims Jr. (1999), the SuperLeader--one who seeks to develop leadership skills in all members of the organization--must provide orientation, guidance, and direction. Super Leaders provide meaningful direction by understanding individual needs and creating a group of people with a shared vision. The source of wisdom and direction still rests with the leader (Manz & Sims Jr., 1999). Providing the source of direction in a scholastic environment, where the variables are constantly changing, is the challenge that leaders face. What can school leaders use to direct decisions and provide the consistency that product-based businesses use? Can schools implement an industrial-based management system to improve school performance? Educational goals can be measured through the use of valid and credible data. Data can be used to
demonstrate a student’s ability to perform on tests during high school, as a predictor for his or her ability to succeed in post secondary facilities, and to indicate a student’s career path. Leaders would be wise to utilize data to evaluate programs, teaching strategies, curriculum delivery systems, and the success or failure of almost any educational effort. The advancement of technology has created readily available data allowing school leaders to gain a clear understanding of student performance, and direction for school-wide consensus decision-making. Enabled through the use of data, consensus decision-making can lead school organizations to develop a strong sense of culture. When members of an organization share a view of the world around them, and their place in that world, culture exists (Hoy & Miskel, 2005). Strong cultures have shared values and a consensus on “how we get things done around here” (Hoy & Miskel, 2005, p. 171). Readily available data provides solid ground upon which to make decisions, and to develop the type of Super leadership needed to move schools into the 21st century.

Data-influenced decision-making is a fairly new “buzz word” that is used extensively in many educational journals on the subject of school improvement and student performance studies. Research consistently indicates that school leaders may raise student performance rates if they use relevant information to form educational objectives, from which corrective action plans, are made. There are several agreed upon factors that influence the success of school systems that routinely perform at high-levels on state-mandated tests, in spite of circumstances that might predict otherwise. Most educational literature will specify the principal as the primary leader in the instructional process. If the principal plays such a crucial role in student
and school-wide success, what do successful principals do that sets their practices apart from less successful principals? Do successful principals utilize data-influenced decision-making on a regular basis to guide instruction? Does data-influenced decision-making significantly impact student results on state-mandated tests?

Deciding on what an organization will focus on in order to make interventions easy for teachers to understand, and to clearly identify the goals, is a difficult process for public school systems. Schools are expected to produce students capable of passing state grade-level assessments, college and trade school entrance exams, students who model good character and students who will become productive members of society. TPS is a quality-based industrial management model that focuses on getting things right the first time by developing leadership skills in all employees at every level of the organization. By creating an environment where continuous improvement is expected from all levels, TPS (Toyota Production System) can view problem solving as a step towards excellence. Many of the procedures used by TPS, are foreign to the educational setting; however, show promising relationships that may be beneficial to service related organizations, such as schools. The researcher will examine existing data to determine the impact that (TPS) makes on student assessments and teacher efficacy at the fifth-grade level. Can TPS, utilizing data-influenced decision-making and systematic procedural changes, lead students to greater levels of performance?

When thinking about schools for the 21st century, two fundamental characteristics come to mind: learning is contextual; and school is a process, not a place (Thornburg, 2002). School leaders make decisions every day that will affect the
lives of students, parents, teachers, and communities. What do school leaders use to
guide their decisions in order to assure that student achievement and student success
are always a priorities? The validity of assumptions that have remained unchallenged
for generations needs to be examined (Thornburg, 2002). Super Leaders of the 21st
century will be held to a much higher standard than school leaders of the past;
decisions guiding educational practices must be based on solid data to provide
consistent results. The focus of this retrospective case study is to determine if TPS
will improve student achievement in two fifth-grade elementary math classrooms and
initiate significant changes in teachers’ perceptions of self-efficacy for those fifth-
grade teachers trained or exposed to TPS principles. This research will also describe
the implementation process using the experiences of the researcher, classroom teacher
and the TPS implementation consultant to further clarify findings and offer more
insight into the application of TPS to education.

Research Questions

1. What is the relationship between the introduction of TPS and student math
   academic performance as indicated by student results from the 2006 and 2007
   PSSA state exams and 4Sight predictive benchmark assessments taken during
   the 2006–2007 school year?

2. What is the relationship between the introduction of TPS and teacher efficacy
   of fourth and fifth-grade teachers as indicated by the Ohio State Teacher
   Efficacy Scale during the 2006-2007 school year?
3. What is the relationship between the introduction of TPS with teacher efficacy and student performance as indicated by an interview with the TPS classroom teacher at the conclusion of the 2006-2007 school year?

*Research Hypothesis*

TPS will show significant statistical improvement in the implementation group’s math performance on the PSSA test and 4Sight predictive benchmark assessments, as opposed to classrooms not implementing TPS.

*Identification of Variables*

The variables are:

*Independent variables.*

1. Design of the instruction to follow TPS process flow.
2. Teacher involvement in the TPS model.

*Dependent variables.*

1. Change in PSSA math scores as indicated by individual student test results from the 2005-2006 and the 2006-2007 school years.
2. Changes in group mean performance scores as indicated by numerical scores from 4Sight predictive benchmark assessments for students in TPS classrooms, as opposed to the remainder of the fifth-grade student body during the 2006-2007 school year.

*Conditions*

1. Student participation in state-mandated testing at grades four and five for all students in the test and control group.
2. Participation in five 4Sight tests throughout the fifth-grade year for all students in the fifth-grade.

3. Application of TPS principles through the employment of a TPS implementation consultant.

Significance of the Problem

In the past decade, the use of academic standards has grown from being the foundation for educational reform in about a dozen states, to being the foundation for curriculum and assessment for all 50 states (Reeves, 2004). The language of accountability varies widely from state to state, with words such as indicators, objectives, benchmarks, and standards, all referring to what students should know and be able to do (Reeves, 2004). Standards are not new in the educational process, but the increased level of state testing and accountability are. Schools are held accountable to a state testing system that measures the extent to which students have met the standards of proficiency in certain curriculum areas. In an ideal standards-based environment, students, teachers, and parents know immediately when success has been achieved by analyzing relevant data. If success has not been achieved, then there is no need to wait for an external judgment to be rendered; the difference between the student’s performance and the expected standard can be immediately determined (Reeves, 2004).

Today’s society relies on testing to tell whether or not water is safe to drink, cholesterol is too high, or the dishwasher on sale is the best value. These tests help to ensure public and private safety and individual good health, and, at the same time, provide guidelines for wise financial and economic decisions (Gandal & McGiffert,
Testing in education is thought to hold the same power to bring about the results parents want, which is to have students who are academically healthy. NCLB is designed to ensure that all students achieve at a minimum standard level or above, in order for schools to be considered successful.

Two fifth-grade math classrooms at Bedford Elementary School, located in a rural district in south-central Pennsylvania, will be the implementation sites for TPS. Bedford Elementary houses approximately 810 students in the facility from kindergarten to grade five, 141 of which are fifth-grade students. There will be two classes of fifth-grade students taught by the same instructor (a total of 45 students) participating in the TPS implementation model design. Goals for student achievement will be set at the 85% range for all assessments including: PSSA, 4Sight and local teacher-generated tests. TPS process flow is designed to surface and solve problems at their root cause by utilizing data to influence instruction. “Despite persistent problems of quality, equity, and scale, many Americans seem to believe that work in education requires common sense more than it does the sort of disciplined knowledge and skill that enable work in other fields” (Ball & Forzani, 2007). TPS process flow is designed to utilize the scientific method to determine success or failure of implementation.

School leaders must make use of relevant data to guide decision-making in public schools. Assessment information must be derived from reliable resources that clarify what they measure, as well as measure what society values. When looking at the data carefully, leaders can make decisions based on facts. The assessment is over and the results have come back. Now, with all the information in hand, the teacher
The Toyota Way can offer a prescriptive remediation scheme to resolve the failure to succeed (Renter & Hamilton, 2003). Assessments provide information, or data, regarding areas where students have succeeded or not. If data does not influence decision-making in schools, there can be no accurate diagnosis of instruction or curriculum weaknesses and no focused remediation efforts. School systems have not relied heavily on data to make decisions in the past; successful schools of tomorrow will be forced to make systemic changes that endorse and mandate the use of data and reliable management strategies to focus change efforts.

During this retrospective case study, PSSA scores and 4Sight Predictive Benchmark Assessment scores will be analyzed to determine if TPS had a significant impact on student performance. The sample will be fifth grade students at Bedford Elementary School in Bedford County, Pennsylvania that have taken the PSSA exam in both the 2005-2006 and 2006-2007 school years. There are 45 students in two separate math classes, each of which receive instruction through the TPS model and are instructed by the same teacher. The remainder of the fifth grade group totals 96 students and will make up the control group for statistical analysis. Mean score variations will be examined using a repeated measures, one between/one within design, to determine if TPS instruction design had a significant effect on student performance. Teacher efficacy will be measured, using the OSTES, to determine if there are gains in perceptions of efficacy between fifth-grade teachers and fourth-grade teachers.

There is much to be learned about the application of TPS to other organizations outside of the car-manufacturing world. Concerns for implementing
TPS in other organizations have ranged from organizational culture to a lack of leadership (Balle, et al. 2006). When managers and organizational leaders fail to understand the framework that guides TPS, they often miss the point of the tools, and therefore, fail to gain the expected results (Balle, et al. 2006). TPS is more about improving organizational performance, identifying problems, solving them the right way, and in doing so, continually increasing the intellectual capacity and skill of all members of the organization (Balle, et.al. 2006). In addition, this study will seek to clarify findings by telling the implementation story with a descriptive analysis of the study through the eyes of the researcher, the implementing teacher and the TPS consultant.

To further describe the sample and limit threats to the validity of results, a more detailed study of the sample is necessary. Within the TPS sample, comprised of 45 students, there are 23 male students and 22 female students. The remainder of the fifth-grade class is comprised of 96 students, of which 48 are males and 48 are females. Using the free and reduced lunch program as a guide for students who may be at an economic disadvantage, cafeteria staff was able to break down the number of students in both the TPS group and the control group for closer examination. Providing only the number of students in each group, with no names attached, the TPS group has 15 students receiving free or reduced lunches, while the control group has a total of 40 students in the program. Of the 45 students in the TPS group, 15 students, or 33%, are in the free or reduced lunch program. The control group is made up of 96 students, 40 students, or 41% are in the free or reduced lunch program. The eight special education students in the TPS group account for 18% of the test group.
The control group comprised of 96 students contains 15 special education students, or 16% of the sample. As shown in these data, the TPS group and the control group are reasonably well-balanced, with only a slight variation in economically disadvantaged students, more of which are in the control group, and special education students make up 2% more of the total population in the TPS group than in the control group. The Bedford Area School District, for the 2006 end of the year reports, shows only 1% of its total population being of minority status. Closer examination of the data reveals that no single category of minority status is comprised of even 1% of the total population in the district, making this statistic immeasurable.

Definitions

Adequate Yearly Progress - Minimum achievement levels that students are expected to meet under the federal No Child Left Behind Act (Renter & Hamilton, 2003).

Data - Information gained from reliable sources that can be used to further understand performance levels (Hoy & Miskel, 2005).

Data-Influenced Decision-Making - Utilizing data to affect decisions within an organization (Hoy & Miskel, 2005).

Hansei - Continuous reflection process aimed at producing employees who are self-learners (Liker, 2004).

Jidoka - Automation with a human touch, building in quality as one works (Liker, 2004).

Just-in-time - Production method designed to deliver exactly what resources are needed and only when they are needed (Ohno, 1988).
Kaizen - Continuous improvement design incorporated into the Toyota Production System (Liker, 2004).

Modus Operandi – A distinct pattern or method of operation (Wikipedia, 2006).

No Child Left Behind (NCLB) - Federal legislation designed to hold schools accountable for student performance (Renter & Hamilton, 2003).

PSSA - Pennsylvania System of School Assessments designed to assess student knowledge on specific predetermined state academic standards (Popham, 2003).

Pull System – Process by which information is derived from customer demand to determine production output in the Toyota Production System (Ohno, 1988).

RAND - Research Corporation for policy-oriented summaries of individual published, peer-reviewed documents or of a body of published work (RAND Education, 2004).

Self-Leadership - An employee’s ability to develop initiative within an organization, following the goals and expectations of that organization (Kelley, 1999).

4Sight Benchmark Assessments - Student assessments, developed by the Success For All Foundation in collaboration with Center for Data-Driven Reform in Education (CDDRE) and endorsed by the Pennsylvania Department of Education, used to predict success on the state’s PSSA exam (Success for all Foundation, 2005).
Standards - What students should know and be able to do (Popham, 2003).

Standardized Tests – State-designed tests used to measure student and school success (Popham, 2003).

Super leadership - Leadership style that focuses on developing all organizational members to be self-leaders (Manz & Sims, 1999).

Toyota Production System (TPS) - An industrial-based management system for a total organization that uses the scientific method to meet customer need (Liker, 2004).

Assumptions

It is assumed that:

Students in the TPS implementation classrooms are heterogeneously grouped, similar to other fifth-grade classrooms in the district. The teacher selected by the district to implement TPS was chosen through a volunteer process, but had to be a grade five math instructor at Bedford Elementary. There are no modifications made in the scheduling process or in the grouping of students to accommodate the TPS classroom. Fifth-grade students in the TPS model classroom follow the same curriculum and academic structure as all other students in the fifth-grade at Bedford Elementary. Students took the PSSA mathematics tests in grade four during their 2005-2006 school year and will be taking the PSSA test again in grade five during the 2006-2007 school year.

Students in the test group will receive instruction through a TPS model according to individual need as indicated by data including, but not limited to, the PSSA, 4Sight benchmark assessments and other local sources of data such as:
teacher-generated tests, observation, assignments and activities. One instructor, the
original classroom teacher, will teach a total of 45 students, comprising two fifth-
grade math classes, using the TPS model. TPS will be applied in the two-targeted
classrooms with the help of an established TPS implementation consultant. Select
administrators and teachers will be included on an established help line; to be used
when immediate assistance is needed.

Limitations

Academic rigor may vary from teacher to teacher and from class to class.
Student academic achievement may differ as a result of other influences not
identified, such as: budget expenditures, teacher preparation time and the availability
of personnel resources to the instructor. Other factors may include differences in the
model TPS classroom’s student population compared to the control group’s
population and the small sample size for teacher efficacy. The impact of these
limitations will be further discussed in Chapter Three.
CHAPTER TWO

REVIEW OF LITERATURE

For too long, education data were too slow in turnaround, too unwieldy to manage and too disparate to make meaningful comparisons. Today, recent advances in computing and communications technology have made possible and practical the widespread use of data for decision-making at every level of the system, from students and teachers, to parents, administrators and community partners (Killion & Bellamy, 2004). If data results can be simplified, everyone involved in the educational process can assess how well they are doing, and modify or continue their behavior accordingly. Teachers find the use of data results extremely helpful in enabling the understanding of individual student learning.

In the following literature review, both the benefits and disadvantages of using data to influence decision-making in public schools will be discussed and reviewed. It is important to note the many differing views on the use of data to influence decisions, as well as the validity of any standardized test to provide the data for such decisions to be made. The RAND Corporation, a nonprofit research organization, released a PDF document in 2004 to address improving educational accountability through lessons from other sectors. The document states that in 2001, with the U.S. Congress approving NCLB, accountability in education sustained enormous gains that are here to stay. NCLB is a performance-improvement effort based on real student outcomes and consequences for school systems if they fail. Failure is no longer an option for public schools. The Rand Corporation’s research examines the
The Toyota Way

educational system to see if management systems from other sectors may be incorporated into education. This review will discuss recent studies and literature regarding the application of TPS, and its use of data to influence decisions in sectors outside the industrial world.

Necessity of Data

Student success is the ultimate goal of school reform laws, policies, and procedures for implementation. With the advancements made in the technological realm of education, educators can now access data from many different sources. Most school reform efforts today focus on student achievement results to guide student performance improvement efforts and to make better school-wide decisions. The educational leader’s capacity to make good decisions rests on his/her ability to obtain information that is accurate and relative to the problem being addressed. Educators can utilize student test results to provide: information upon which students need to focus, direction for school improvement efforts, and long term documentation about school performance. Test results not only provide school-wide data, but also offer a comparison of similar schools across the state to further enhance positive decision-making. Data-influenced decision-making is a continuous process of monitoring and adjusting efforts to address areas of focus as indicated by the data (Keefe & Kelley, 1990).

Data will indicate learning strengths and weaknesses, and specify what sub-groups of the student population are having difficulties, and in what areas. Both instructional changes and professional development can be predicted through the use of data. Attention may be placed specifically on materials and equipment used, which
can be monitored for its impact on student achievement. Keefe and Kelley (1990), describe effective school improvement in eight steps. The first step involves a careful review of district philosophy and goals. Next, data must be collected to affirm the goals, which should be community-based. Goals must be operational and valid measures should be selected. In the first three steps of the school improvement process, Keefe and Kelley (1990) identify the use of data and accurate measures of success as being essential to school improvement. A great deal of information is available to educators in the form of data; however, the question regarding whether or not the data being utilized are accurate and measurable through standardized tests still remains.

One building level question asks if actions are being taken because they are the way things have always been done, or because there is tangible evidence to support them (Johnson, 1997). Currently there is an enormous amount of data available to school systems; the effective use of these data is the only way to ensure that it will help student performance. According to Johnson (1997), archival data can be used to set baselines against which future performance can be judged. Analysis of data use should be based on a continuous model. Data should not be interpreted for punitive purposes, or on a grade-level “static” basis. All students are unique; causing variable responses to different strategies, making a one size fits all structure very problematic. Teachers may discover their best approach to success in instruction through studying their students on an individual basis. Data interpretation should be used to gauge the achievement of clearly defined goals and plot a course for future performance (Johnson, 1997). According to Gabel, Arllen, and Whinnery (1997), by
using curriculum-based assessment, graphing the results, and interpreting the data, student performance and group performance patterns can be identified. Plans can then be made to address individual needs based on a careful analysis of the data. If further analysis leads to a decision to introduce an intervention, direct assessment of that intervention should take place (Gabel, et. al, 1997). To better understand school achievement, the school must learn to utilize relevant data (Holcomb, 2001).

Useful educational assessments must make clear what is being measured and measure what we value most (Gandal & McGiffert, 2003). Assessment data provide valuable information on where students and schools need to improve. Educators get the most useful information about students when they compile data from a number of sources, including classroom assignments, quizzes, diagnostic tests and large-scale assessments. Together, these tools paint a more detailed picture of student performance than a single assessment (Gandal & McGiffert, 2003).

Educators often begin improving their schools by asking two questions: “What data should we be analyzing to help our school improve?” and “What data other than state standardized test results can we use?” (Bernhardt, 2003, ¶ 1). Sorting out the data needed for school improvement is a task for educators, along with finding a tool, or computer program, to make the data easily accessible. According to Bernhardt (2003), almost any question about the effectiveness of a school can be answered by gathering, intersecting and analyzing four kinds of data; demographic data, quantitative data, perceptual data, and school process data. Schools will need access to various types of demographic data that describe the students, staff, school and surrounding community. This information delineates the context in which the
quantitative data are essential and should consist of a variety of measurements including: norm-referenced tests, criterion-referenced tests, standards assessments, teacher-assigned grades and authentic assessments. These types of data should indicate the impact of a school’s education delivery system on student performance. Perceptual data gathered through questionnaires, interviews and observations help schools understand what students, parents, teachers and the community believe about the learning environment (Bernhardt, 2003). School process data are also needed to analyze school programs, instructional strategies, assessment strategies and classroom practices.

Looking at the intersection of two kinds of data over time allows schools to examine trends that may be developing. For example, standardized achievement test scores disaggregated by ethnicity over the past three years can help a school see whether the scores of a given ethnic group, compared with those of others, constitute a trend or just a fluctuation (Bernhardt, 2003). Intersecting three of the categories of data generates a comparison between a given group’s achievement scores and the same group’s questionnaire results. Observing these data sets over a period of time allows educators to view trends, to understand the learning environment from the students’ perspectives and to know how to deliver instruction in order to get the best possible results for all students (Bernhardt, 2003). Data allows schools to determine the areas in the school program that least contribute to student success and areas that prove to be most productive. This provides a sound basis for staff development programs, curriculum revisions and instructional strategies that work best at various
ages and grade-levels. Under NCLB, schools need to show advancement to avoid being identified as needing improvement. To avoid being so identified, a school must not only show progress in overall test scores, but also must improve achievement for every subgroup of students in the school, including racial and ethnic minorities, special education students and students with limited English proficiency (Rentner & Hamilton, 2003).

Today’s educators have a public mandate to make instructional decisions on comprehensive, accurate data (Perkins, 2003). Taxpayers are demanding accountability of school systems across the state and are looking for proof in the form of quantitative data. The Education Commission of the States report that 41 states require that schools release student assessment scores to the public and many states require the reporting of additional indicators, including drop-out rates, student attendance and expenditures (Education Commission of the States, 2000). Comparative data across schools and districts make it easier to differentiate the practices and policies that work from those that do not. Savvy educators across the nation are embracing performance data as a useful means for directing school improvement (Killion & Bellamy, 2004). Fear and mistrust of data are giving way to a new culture of use in which teachers and administrators routinely collect and analyze student data to achieve goals (Killion & Bellamy, 2004).

Concerns Regarding Data Use

Even with all the overwhelming evidence that data-based decision-making is a necessity for schools to survive in the 21st century, many educators believe that standards-based testing and data-influenced decision-making are misleading through
faulty information and tests that do not measure what is essential. Many people also believe that such data-based decisions are undermining a teacher’s ability to expand upon student interests because demands for covering testing material are great.

Data is an appealing word to individuals seeking improvement and accountability standards for schools because it measures actual student performance on pre-determined standards. According to some educators, data can be confusing and can actually suppress the educational process if it is inaccurate or not used properly. According to Popham (2003), most state accountability systems fail to produce the kinds of data that will improve teaching and learning. Popham (2003) goes on to state that he wants educators to realize that the wrong kinds of data, even if warmly applauded by many, can actually stifle teachers’ pursuit of accurate evidence regarding their students’ performance. Teachers are constantly reminded of test scores and how their students performed on various tests. Educators are even encouraged to make decisions based on such data. To avoid becoming disillusioned with all data, teachers must learn how to distinguish between “instructionally delightful” and “instructionally dismal” data (Popham, 2003).

According to Popham (2003), for a test to be useful in instruction, it must meet five categories. The five categories are as follows: significance, teachability, non-intrusiveness, describeability and reportability. Significance in a test measures students’ attainment of worthwhile curricular content. A test can provide no useful data if it does not measure the material of importance. Teachability is when a test is used to measure something that is teachable. Teachability enables most teachers to have students master test-aligned material (Popham, 2003). Non-intrusiveness refers
to the actual time that the testing process takes away from instruction. Well-designed tests will use a minimal amount of time to measure the desired outcomes.

Describability in a test is based on clear descriptions of the skills and knowledge it measures. Reportability indicates whether or not test results are specific enough to inform teachers about the effectiveness of the instruction they provide. According to Popham (2003), today’s standardized tests do not meet up to these expectations for useful data results.

Decisions affecting the classroom are made at the state, district, and building levels. Often the large-scale state tests are overly emphasized in the decision-making process. This disenfranchises teachers who believe that classroom assessment/evaluation is ignored (Harris & Carr, 2001). In a study that measured achievement gains at a Chicago public school in 1994, results indicated that much of the data gained from norm-referenced test were not applicable in reform efforts. The use of status indicators, especially in urban areas, where certain demographic characteristics will likely obscure causal interpretations is questioned. According to Bryk and Deabster (1994), the following problems are associated with norm-referenced tests: growth among non linked tests can be misleading, tests need to be linked-equated across grades in looking at change over time, and tough retention policies make status comparisons problematic. High-stakes test results are only one of many sources of data for making important decisions in instruction (Love, 2004).

“You don’t have to go far or look hard to find data being abused” (Love, 2004, ¶ 1). Abuses of data often include single and imperfect measures for performance levels of students. High-stakes testing is many times the culprit of data
abuse as these tests frequently attempt to evaluate school systems and create further accountability. High-stakes standardized tests are often used to size up the effectiveness of schools, dole out financial rewards and punishments and determine students’ futures (Love, 2004). According to Love (2004, ¶ 2), “It is a big problem because many high-stakes tests themselves are seriously flawed.”

Application of Data

In a study of the Illinois school system’s approach of utilizing data-driven decision-making to improve school learning environments and promote pro-social behavior, data results showed an increase in student attendance and positive behavior. Data collected over a two-year period indicated that schools utilizing the Positive Behavior Interventions and Supports (PBIS) approach and data-driven decision-making sustained lower rates of suspensions and office referrals (Atkins, et al., 2002). It is unclear exactly what the impact of higher attendance rates and lower disciplinary referrals had on student achievement, but it was suggested that student performance should improve through increased time-on-task.

The Department of Special Education at The Virginia Commonwealth University conducted a similar study to determine the effects of data-based decision-making to improve learning for students with disabilities. Project ALIGN (Supporting Data-based Decision-Making to Align the Intent and Implementation of IDEA with the Goals of the National Education Reform) analyzed: funding formulas, transition requirements, roles of economic and demographic factors and the diverse background of students to determine the effect these factors have on student placement and student achievement. Through a series of interviews and comprehensive data analysis,
project ALIGN indicated that a strong correlation existed between economic and demographic factors, and student rates of identification, placement and graduation (Oswald & Coutinho, 1997). Summative evaluation of data demonstrated the need to analyze national, state and local data to make decisions regarding placement and education strategies for special needs students (Oswald & Coutinho, 1997). Research supports the concept of data-influenced decision-making to improve student outcomes in both academics and behavioral support. Finding appropriate data that are worthy of affecting instructional practices--that is relevant, grade-level appropriate data that actually test what is important--is a constant challenge for educators.

In a study examining the overlap of content in four standardized tests and five major reading series, Shapiro (1987) found that discrepancies existed between the different standardized test results and the content of the five major reading series. “A particularly interesting result of the present study is that overlap diminished as grade-level increased” (Shapiro, 1987, p. 64). In general, these findings suggest that cautions should be used in the interpretation of results from reading sections of the achievement tests (Shapiro, 1987). School systems are held accountable to state-standardized testing instruments that may not measure actual student learning when compared to the curriculum being taught. Further studies may indicate that high-stakes testing has little effect on student performance when compared to low-stakes tests results.

Greene (2003) proposes that high-stakes tests can accurately measure student proficiency, and when compared to the results of low stakes tests, there is little variance in the student results. In Greene’s study, a test was considered to be high-
stake if it affected graduation, promotion, accreditation, funding cuts, teacher bonuses, school grading/ranking systems, or state assumptions. High correlations existed between the high- and low-stakes tests outcomes, indicating that the stakes attached did not alter outcomes (Greene, 2003). The study points out that standardized tests can measure student proficiency, according to the content of the exam, but they may not measure student learning as it relates to the taught curriculum. According to Greene (2003), carefully designed low-stakes tests can provide data to make instructional decisions, as low-stakes tests tend to yield similar student results from that of high stakes tests.

Other Factors Affecting Student Success

Often overlooked by decision makers are the more subtle factors that influence student performance. In a study designed to examine how instructional time and learning rates affected student performance, researchers found that both factors need to be considered when making instructional decisions (Cates, et al., 2003). The results clearly indicate that decision makers should involve the school psychologist more often in the decision-making process, because students learn at different rates. By taking into account the amount of time required by individual students for various interventions, educators can enhance data-influenced decision-making to benefit all students.

Protheroe (2001) indicates that data-based decision-making is an important technique that school systems should employ to improve student outcomes. Finding and using data that are meaningful to instruction is a necessary strategy for success rates to improve. Educators must learn to ask the right questions (Protheroe, 2001):
What should students know, and how should they use what they know? How well should students perform? How will student performance be assessed? How well do students actually perform? What will be done to improve student performance?

“Many districts that have made effective use of assessment data found early on that what was needed as a first step was an intensive review of their curriculum” (Protheroe, 2001, ¶ 1). Schools must begin to view testing not as a means for judging students, but as a tool for improving learning.

According to Protheroe (2001), good data can be easily disaggregated not only by district, but also by classroom and specific groups of students; good data will provide a detailed analysis of results by objective, or skill, in addition to overall scores. “The data not only helps [sic] teachers see specific areas of difficulty for each student, it also helps teachers and principals to pinpoint objectives that either need to be covered more thoroughly, or taught in a different way” (Protheroe, 2001, ¶ 3).

“Educators across the country who have learned how to effectively use assessment data have indeed ignited change and achieved positive results at the district, building, classroom and student levels” (Protheroe, 2001, ¶ 2).

In 1993, the Massachusetts public school systems implemented a reform effort that focused on high standards and high-stakes for all (Reville, 2004). This reform effort grew out of local concern that students were not being well prepared by the public school system, and that students would not be ready for future expectations under the current system. During the reform efforts that began in 1993 and ended in 2003, a detailed study took place to determine the effect that high standards and high-stakes had on student achievement. “Massachusetts’ standards, while difficult and
time-consuming to develop, have been rated by organizations such as Achieve, Inc., as some of the best and highest in the nation” (Reville, 2004, p. 593). The results of this long-term study indicate that students did indeed perform at higher levels as a result of the concentrated statewide effort. On the 2000 National Assessment of Educational Progress (NAEP) science exam, 43% of fifth-grade students scored at or above proficient, which was the highest in the nation (Reville, 2004). On the 2003 NAEP reading test, fourth-graders tied for first nationally, 43% of the eighth-graders scored proficient or above, and high school students outscored the national averages (Reville, 2004). The first graduating class exited the school system in 2003 with 95% of all students passing both the math and reading exams. “The challenge ahead is to hold everyone in education responsible for providing the teaching and learning conditions that will enable all students to attain high standards” (Reville, 2004, p. 596).

Some educators have come to believe that high-stakes tests lead to lower levels of problem-solving ability, especially when schools teach to the test. Spending too much time on test-taking skills led to memorization and a decline in students’ abilities to solve uncommon problems (Cankoy & Tut, 2005). A definite correlation existed between the amount of time spent preparing for standardized tests and student results. Decision-makers must be careful not to place too much emphasis on tests, as student and teacher creativity are restricted, thus limiting student problem-solving ability (Cankoy & Tut, 2005).

Other factors, such as staff development, fiscal allocations and staffing patterns must be taken into account when studying the factors affecting student
performance in high-achieving school systems. Huffman (2003) concluded that, in a study of professional development activities and teacher participation, most staff development activities never reach the classroom for implementation. Five types of staff development were analyzed to determine to what extent teachers’ implemented learned practices into their classroom. Three of the five processes—immersion, curriculum implementation and collaborative work, proved to be ineffective for teachers. “The results suggest that, for science and mathematics teachers, participation in two types of professional development, namely, examining practice and curriculum development, are related to the use of standards-based instructional practice” (Huffman, 2003, ¶ 1).

Fiscal and staffing allocations are key factors that can limit a school system’s ability to effectively manage change and improve student performance. In a two-year study that focused specifically on spending patterns and school performance, Bray (2003) concluded that improvement districts showed a range of reforms to boost student achievement and allocations of resources to support their efforts. Districts that were identified as high performing showed greater spending on instruction and lower spending on general administration and no instructional services than did their comparison districts (Bray, 2003).

As previously stated, student achievement is the ultimate goal of school reform laws and the rules, policies and procedures for implementing them. Many federal and state laws require schools to implement improvement plans and to set goals to increase student achievement of standards. Goals for improvement are based on assessment results and the indicator systems built within the accountability system.
These results reveal overall learning, conditions that affect learning and discrepancies in learning between content areas, groups of students and grade-levels. Using this information, school leaders can determine what should be done to improve and how improvement might be accomplished to increase student achievement.

As indicated earlier in this literature review, schools that learn to utilize data to influence instruction and make decisions regarding instruction, typically perform at higher levels. Daggett writes in an article, for the International Center for Leadership in Education, that there are nine common characteristics used to raise student achievement; one of those is a laser-like focus on data (Daggett, 2004). The laser-like focus utilizes data from all groups of students in a school system. Those data are then applied to the decision-making process to guide instruction. Data analysis provides focus for administrators to target areas of strength, as well as to identify areas of weakness within the instructional process.

During this study, the researcher will analyze existing data to determine the effects of a manufacturing-based management system incorporated into the public school setting, for the purpose of implementing data-influenced decision-making and procedural changes, as it relates to student performance and teacher efficacy. TPS will be the industrial-based management system implemented into a public school setting at the fifth-grade-level.

*Toyota Production System*

TPS is the most recent and one of the most obvious efficient production systems to evolve since the mass production system invented by Henry Ford in the early Twentieth Century. Mass production was noted for its high rates of production
per individual worker, producing products more cost-effectively and with greater consistency. At the height of Henry Ford’s operation, the factory employed over 16,000 people. Changes in economy, fuel costs, material costs and customer demands for more versatility in design eventually necessitated change in the manufacturing operation, so that companies could continue to compete in a global society. While many of the automobile manufacturing plants continued to utilize the mass production process, Toyota was making adaptations and changing its processes as the demands placed upon it changed. Toyota has never mass-produced products in a traditional way (Sobek & Jimmerson, 2001).

Toyota grew from a family business started by one man, Kiichiro Toyoda, in 1938. Today the company employs 240,000 people from around the world. The philosophy that leads the Toyota system has remained remarkably consistent from 1938 to the present day. Toyoda was a tinkerer and an inventor who grew up in a rural farming area in Nagoya. He grew up in a family that struggled for financial stability and his parents worked very hard weaving for a living. His initial invention grew from his desire to see his family members progress and not have to perform manual labor. In 1894, he developed a wooden spinning machine that was cheaper and worked better than any existing machine at that time. Not long after his initial invention, Toyoda decided that the machine should be powered by something to speed production and eliminate the need for people to spin the yarn wheels. At that time the only source of power being utilized was the steam engine. In 1926, he introduced Toyoda Automatic Loom Works, the parent organization of the Toyota Group. The TPS philosophy, still in existence today, originated from this parent
company back in 1926. “Toyoda figured out how to make things work by trial and error, and getting his hands dirty. This approach would become part of the foundation of the Toyota Way” (Liker, 2004, p. 16).

Toyoda eventually went on to design an automated and sophisticated power loom that became famous. One of his major inventions with the power loom was the development of a device to stop the loom whenever a thread broke. This invention evolved into one of two pillars of the TPS and is referred to as “jidoka,” which means automation with a human touch (Liker, 2004). This technique allows for quality to be built in as materials are produced. It also frees up workers so they are not tied to machines and can perform more valuable work. TPS is a unique system in that it encourages all workers to be part of the problem-solving environment through a culture of continuous improvement.

Toyota is a remarkable success story of a business that originated with a simple wooden yarn machine and evolved into one of the three largest automobile manufacturers in the world. Toyota’s annual profit for 2003 was $8.13 billion, a larger profit than the combined earning of Ford, General Motors and Chrysler (Liker, 2004). During a time when other automobile manufacturers were suffering, Toyota was able to earn a 24% increase in stock shares, with a total market value of $105 billion for the company. Toyota has long been regarded as the leading auto maker in Japan, but in 2003 Toyota outsold Chrysler to become the third largest auto manufacturer in the world, leaving only General Motors and Ford ahead in sales. American companies are taking two to three years to design a new product, while Toyota is capable of taking a vehicle from design to production in one year. The
Toyota Camry was the top-selling passenger car in the United States from 1998-2003, outselling all American made passenger cars in its class. The Lexus was introduced by Toyota in 1989 and outsold BMW, Cadillac and Mercedes-Benz from 1999-2002. “Toyota is benchmarked as the best in its class by all of its peers and competitors throughout the world for high quality, high productivity, manufacturing speed and flexibility” (Liker, 2004, p. 5). J.D. Power and Associates and Consumer Reports magazines have consistently ranked Toyota at the top of quality products.

Much of this growth came as a result of the oil crisis in 1973, which was followed by a recession. Japan’s economy had nearly collapsed, but in the midst of all that was happening, Toyota Motor Company sustained greater earnings than any other company in 1975, 1976 and 1977 (Ohno, 1988). The two pillars that pulled Toyota through the recession and eventually helped to form the company that sustained higher profits than any other automobile manufacturer are: just-in-time and jidoka. By supplying parts as they are needed and when they are needed, Toyota can minimize wasted materials and idle workers. Toyota began to look at production in a completely opposite way than American companies did. Toyota invented “lean production” (also known as the Toyota Production System or TPS), which has over the last decade triggered a global transformation in virtually every industry to Toyota’s manufacturing supply chain philosophy and methods (Liker, 2004).

Womak and Jones (1996), in their book titled, Lean Thinking, describe lean manufacturing as a five-step process: defining customer value, defining the value stream, making it “flow,” pull, and striving for excellence. “All we are doing is looking at the time line from the moment the customer gives us an order to the point
when we collect the cash and we are reducing that time line by removing the non value-added [sic] wastes” (Ohno, 1988, p. ix). TPS focuses on the big picture and suggests that often the best thing to do is idle a machine and stop producing parts until a solution to a particular problem is found. Sometimes it is not the best solution to keep workers in production every minute of the day. Only produce at a rate the customers’ needs demand. TPS research also suggests that automation is not always the best solution, even when it can save money and time. The most flexible resource a company has is its people. Every move that is made by TPS starts with valuing customer perspectives.

Toyota has specifically identified eight types of nonvalue-added waste that have a negative effect on organizations: overproduction, waiting, unnecessary transport or conveyance, over processing or incorrect processing, excess inventory, unnecessary movement, defects and unused employee creativity. Overproduction refers to producing items for which there are no orders. Waiting means workers are idle, merely watching a machine produce parts. Unnecessary transport means carrying work long distances, or inefficient operations causing parts to be moved in and out of storage. Over-processing is simply taking redundant steps to produce items due to product design, poor tools, or other obstacles that cause interference to the process; it is doing more than the customer asks for. Excess inventory causes longer lead times and results in damaged goods because of things being moved around too many times. Unnecessary movement is any wasted motion made by employees that does not result in better customer service, or increased productivity. Defects require employees to rework and fix items, resulting in wasted time and parts. Possibly the most critical
The Toyota Way

factor, according to TPS, is unused employee creativity. Toyota defines this non-value as: losing time, ideas, skills, and improvements and learning opportunities by not engaging or listening to employees (Liker, 2004). TPS supports the idea that pushing employees to produce when there is no need according to customer demand reduces employee motivation to continually improve operations.

Toyota operates from 14 basic principles that support the management system and philosophy that guides their company in the business world. Toyota believes that organizations should base their decisions on long-term philosophy, even at the expense of short-term financial goals. “Work, grow and align the whole organization toward a common purpose that is bigger than making money” (Liker, 2004, p. 37). Organizations should create continuous process flow in order to surface problems. Unlike many organizations, Toyota sees problems being brought to the surface as growth. Production and labor decisions should be based on customer demand. Toyota refers to this process as a pull system, or looking at manufacturing in reverse. The workload should be leveled out and shared equally among all members of the organization (heijunka). Within this effort to share the workload, Toyota states that it wants its employees to “work like the tortoise, not the hare” (Liker, 2004, p. 37). It is important to build a culture that stops to fix problems. Within any organization there should exist a system to quickly solve problems and put in place corrective measures. Toyota uses stable and consistent tasks for its employees so they can build quality into the product through their experiences. Maintaining repeatable methods allows Toyota to better predict performance and quality. Visual inspection is the primary method of quality control. Designs should be simple and reports should be very short.
The Toyota Way

and concise—never more than one written page in length. If technology is to be used, it should be previously tested and proven reliable. Technology should support people—not add to their workload. Toyota believes in growing its leaders from within the organization, developing people who thoroughly understand the system and type of work they perform. Creating a culture of continuous improvement, run by exceptional employees who share a common philosophy, is how Toyota maintains high-levels of employee motivation. “Respect your extended network of partners and suppliers by challenging them and helping them improve” (Liker, 2004, p. 40). Toyota also believes that decisions should be made slowly and by consensus; once a decision is made, the implementation should be rapid. The final step in this process is to become a learning organization through continuous reflection and continuous improvement. At the conclusion of this 14-step process, Liker (2004), states that the company protects its organizational knowledge base with the development of stable personnel, slow promotion, and very careful succession systems.

TPS in Other Sectors

The question that remains is: Is TPS adaptable enough to serve in other sectors of the business and service world? Without question, the TPS effectively does service and fuel a healthy manufacturing-based system in a competitive global marketplace. Despite the fact that the TPS management system has its roots formed in the industrial arena, Toyota leaders believe that it can work in any management system if the philosophy is adopted in its complete form. TPS is grounded in utilizing the strengths of team members and workers, forming a philosophical base serving as a foundation from which to build. In a special release, compiled by the Public Affairs
Division of the Toyota Motor Corporation in 2003, leaders state that TPS will not change in the future and that it will be able to meet any challenge. “An environment where people have to think brings with it wisdom, and this wisdom brings with it kaizen” (Minoura, 2003, p. 2). Education must become more like a business, with improving so-called production. While businesses train their employees to become highly efficient, education seems to lack in the area of employee development. Minoura (2003, p. 5) states, “If we are going to allow Toyota’s DNA to spread and evolve globally, we need to develop and train global people.” In order to progress in a competitive world, businesses need to develop people who are capable of solving problems and implementing new ideas.

Due to public pressure for more cost-effective and higher quality personalized medical care, the health field has become one of the first service-oriented sectors to attempt to adopt TPS. On January 30, 2001 the President of the Community Medical Center (CMC), Grant W. Winn, submitted a proposal to the National Science Foundation seeking support to implement TPS into the medical field at two hospital facilities. The proposal sought financial support to implement TPS into the medical arena in order to reduce costs and regain customer satisfaction. Two questions were of particular focus: Can the principles of the TPS improve healthcare delivery? If so, what implementation strategies are more likely to lead to success (Sobek & Jimmerson, 2001)? Lean is a philosophy, or perspective incorporated into TPS, that is designed to eliminate waste and defects through the pursuit of perfection. CMC identified in its study specific action steps within TPS that developed an understanding of how the culture of “lean” was developed.
Every time an improvement is proposed in TPS, the proposal explicitly states the expected outcome, which can be verified or refuted through experimentation and data (Sobek & Jimmerson, 2001). The US healthcare industry is currently in a state of constant change, some factors causing that change include: rapid technological innovation in equipment, changes in medication, new treatment options, increased training costs, changes in regulations and insurance options. According to Sobek and Jimmerson (2001), healthcare costs are projected to meet and exceed the $2 trillion mark and many customers are dissatisfied with the quality of healthcare. By implementing TPS, CMC is hoping to gain a fast response time from the initial need to treatment of the problem--very similar to the pull method employed by TPS. Hospitals must deliver a product that is free of defects, which is much like the philosophy that Toyota employs in its manufacturing-based system. TPS allows for inefficiencies or negative patient interactions to be addressed in “real time,” because it teaches workers to recognize and immediately correct defects under the direction of a mentor/facilitator/team leader (Sobek & Jimmerson, 2001).

The proposed work in this project is a collaborative effort between CMC and Montana State University (MSU). The pilot studies involved two units at CMC who performed very different job responsibilities. Teams studied the TPS model and its possible use within their assigned unit, and then used TPS tools to redesign their processes. This part of the study was primarily qualitative by using surveys to determine if the changes implemented by each of the teams resulted in more positive customer relations. The second part of the study used qualitative techniques to analyze which TPS strategies are more likely to facilitate success. To answer the
second question the researcher used: direct participant-observation, journaling by team members, artifacts created in the redesign effort and interviews to explore the psychological and cultural barriers that make it difficult for healthcare workers to adopt TPS principles (Sobek & Jimmerson, 2001). The proposal states that full TPS implementation requires a major shift in how employees think about work, even at the lowest levels of an organization.

CMC is a 132-bed facility located in Missoula, Montana. In 2001, CMC employed approximately 1,000 employees and partnered with hundreds of physicians. CMC is a rural hospital that serves a low-income area; it is a prime facility to test the TPS model. CMC also has a working relationship with a neighboring hospital that serves as a provider for the Confederated Salish and Kooteai Tribes of the Flathead Indian Reservation, St. Luke Community Hospital. St. Luke served as a testing area for newly designed training modules and as a partner in this research project (Sobek & Jimmerson, 2001).

TPS was applied to both the Medical/Surgical department and the Cardiac Catheterization Lab at CMC. A multi-disciplinary team consisting of a facilitator, two physicians, two nurses, an administrative representative, industrial engineering faculty, a graduate student and an undergraduate student from MSU were used to introduce the TPS method. During the summer of 2001 the project looked at the cardiology department of CMC for a period of one month to set up the value-stream map that would monitor the flow of patients within the system. TPS principles were introduced to the employees throughout the summer months, and employees were directed to write problem-solving reports. During the 2001-2002 year a seven-week
training course was set up and offered to CMC employees to aide them in implementing TPS ideas in their respective departments. Identical courses and training were also offered at St. James Hospital. During the summer of 2002, the project focused on the pharmacy department to create a value-stream map for pharmacy orders. During this process, TPS uncovered the high number of interruptions faced by the pharmacists that hindered daily operations, even when two pharmacists were on the job. Training sessions were offered to employees throughout the 2002-2003 year, and one-on-one coaching was put in place for value-stream planning. A graduate student from MSU initiated the study and began developing measurement devices during the 2003-2004 year. CMC also began an effort to transition from a dependent organization implementing TPS to an independent, self-sustained system of continuous improvement. In 2004, the initial fieldwork began with survey instruments to measure the effects of TPS on the medical facility of CMC.

Initial research indicates that the first three rules of the TPS, with some refinement, are effective management options for the healthcare field (Ghosh & Sobek II, 2004). The first four TPS rules are: (a) All work shall be highly specified as to content, sequence, timing and outcome. (b) Every customer-supplier must be direct, and there must be an unambiguous yes-or-no way to send requests and receive responses. (c) The pathway for every product and service must be simple and service must be direct. (d) Develop people who are able to solve problems as they occur. These four principles represent the framework for implementation. A cross-case analysis indicates that of the 16 cases examined, high to medium degrees of change
occurred in specification, connection clarity and pathway simplification, with large changes (68-100%) in customer survey feedback. Correlations indicate that the activity specification and pathway simplification are correlated with positive outcomes and customer survey feedback. “The work presented here tests the applicability of a general model for designing lean work processes in the specific context of a hospital. Overall, findings support the model’s applicability to healthcare, leading to both theory advancement and practical application” (Ghosh & Sobek II, 2004, p. 26). The TPS framework provided a systematic method for examining in-depth the work processes of CMC and allowed a much closer examination through activity, connection and pathways. Research revealed that many of the designing work processes were often missing; a finding that likely would not have surfaced otherwise (Ghosh & Sobek II, 2004).

In a similar pilot study conducted by Sobek and Jimmerson (2004), results indicated that TPS helped management to revamp an overburdened pharmacy department by simplifying procedures and having direct pathways. The conclusions indicate that the first four rules in TPS can be applied to healthcare institutions with a high-level of success. The effectiveness of the principles is most likely attributable to the aspect that enables organizations to simplify processes without compromising them. Perhaps the biggest challenge is to move away from the quick-fix mentality that remains prevalent in healthcare, and begin to think about how the system can be improved; TPS principles and tools are designed to help organizations do just that (Sobek & Jimmerson, 2004).
Spear (2005) argues that industry leaders such as Toyota, Alcoa, South West Airlines and Vanguard, have demonstrated that it is possible to manage the contributions of dozens, hundreds, and even thousands of specialists in such a way that their collective effort not only is capable and reliable in the short term, but it also improves steadily in the longer term. Spear states that the problem in healthcare is workers tend to avoid problems when they encounter them, instead of solving the problems immediately. He goes on to state that the system lacks reliable mechanisms to detect and resolve problems from a procedural standpoint. The goal of implementing a system, such as TPS, is to make higher-level employees part of the problem-solving process and not the people who come to the rescue once a problem is surfaced.

The West Penn hospital employed the principals of TPS to resolve an issue in its surgical preparation department (Spear, 2005). On average, prior to the implementation of TPS, one in six patients was scheduled as ready for surgery but did not have appropriate blood work. Patients became frustrated and nurses were left in a frazzled state; doctors waited for preparation at a wasted expense of approximately $300.00 per minute. When the unit reviewed the process, it was able to eliminate unnecessary steps in a systematic way. Visual indicators allowed for a quick visual inspection process through the use of stickers on charts and signs on the ends of beds. Another employee was in charge of taking blood; sometimes it was the nurse and sometimes it was the technician. Assigning a nurse to be solely responsible for taking blood and using a charting system resulted in a sharp decline in the number of patients missing blood tests. The final step in the process was detected when a nurse
realized that he or she did not always know who had been registered and needed blood work. A card system was put into place that offered a quick visual indicator that the patient was registered and ready for blood work. By incorporating these simple, essential systematic changes and refining the waiting areas to improve privacy and comfort, the western Pennsylvania hospital was able to reduce the number of patients waiting for blood work from seven to zero (Spear, 2005). Spear (2005) goes on to state that, “In the highest-performing organizations, all workers need coaching to learn how to reduce ambiguity systematically and to learn how to continually improve processes through quick, iterative experiments” (p.85).

The pharmacy at the University of Pittsburgh Medical Center (UPMC) employed TPS to resolve a problem they were having with filling prescriptions in a timely fashion (Spear, 2005). The pharmacy is responsible for making deliveries of medication throughout the hospital as needed and directed. The process was confusing at best, as nurses would interrupt their work to call the pharmacy, which, in turn, interrupted the pharmacy staff. Doctors would make rounds and write prescriptions all at once, which backlogged the computer system and prevented the pharmacist from correcting any potential hazardous medication issues until it was later discovered when filling the order. To resolve this issue the pharmacy staff spelled out exactly which demands the nursing staff was placing on them were most important. They decided that in order to keep up with the process presently in place, they would have to fill one order every three minutes. After several trials it was determined that they could not keep up under current conditions. Some medications were hard to locate due to storage issues. TPS systems helped them to redesign the
storage system according to need and not alphabetically. Other solutions were more complicated, such as changing timing at which drugs left the pharmacy, the delivery route technicians took in the hospital and the way orders were placed with distributors (Spear, 2005). In addition to the other changes implemented, the pharmacy started filling orders every two hours; the end result was a total reduction in missing medications of 88%. According to Spear (2005), this research demonstrates that TPS can be successfully applied to health care services.

TPS in Education

“Lean Thinking” is presently a topic of discussion for improvement in the educational sector (Marchwinski, 2006). Incorporating an industrial management system that is product-based has been out of the question in the minds of many educators; at least until recent success stories began to surface in other sectors of the service industry. Further study into TPS reveals a structured flow of the system, which continually seeks to surface and solve problems while maintaining a quality product. It is not the manufacturing element that may help other organizations. It is the system of checks and balances that Toyota employs in its management philosophy of the operation that may help education to look at the process in a different and possibly better way. TPS seeks to identify and rapidly solve problems. What are the obstacles that may deter action steps from being implemented into education? Can educators employ an industrial-based management system into the educational process? Can education reduce costs and increase productivity while incorporating TPS? Finally, a bottom line question: Can TPS improve student and teacher performance (Marchwinski, 2006)?
Marchwinski (2006) states the following:

While lean is the prominent focus throughout manufacturing and increasingly all business fields, ranging from healthcare to retail, on the campuses of many colleges and universities, lean hardly appears in undergraduate or graduate curricula and faculty fail or hesitate to teach the principles that business is embracing. That may change as lean-minded professors have joined forces with the Lean Enterprise Institute (LEI) to form the Lean Education Academic Network (LEAN), educators dedicated to implementing and continuously improving lean education in academia. (p.1)

Lean Enterprise Institute in Massachusetts released that news statement to the public on January 17, 2006. The idea is for Ohio State University professors and leaders of “Lean” to identify the current state of lean in academia and develop a vision for the future. While no one seems to know exactly what “Lean” means for education, there is growing interest among educators to find out.

The RAND Corporation looks at several management ideas that may help educational organizations improve student performance. The Malcolm Baldrige National Quality Award is a program that recognizes school systems for performance excellence. The Baldrige award process offers a promising model for schools to follow: processes, resources and data to promote strategic goals. TPS and Lean Manufacturing offer possibilities for education to improve by offering more choices to consumers, more decision-making involvement for workers, and enhanced productivity (RAND, 2004). According to this article, Lean centers around three basic principles: (a) All those involved must understand Value stream. (b) The culture of
an organization must encourage improvement. (c) Customer demand should drive the rate of production. The RAND Corporation defines two areas where they see education making some real gains from TPS. The educational value stream should encourage teachers to see the big picture of a child’s education and understand how their contribution affects the final outcome—which activities are valuable and which activities are wasteful. Creating a culture that makes teachers responsible for hypothesis-testing, experimentation, and continuous improvement could cause an increase in research on educational practices, enhance student outcomes and progress toward the scientifically based teaching practices favored by NCLB (RAND, 2004).

According to Shuell (1992), learners who are engaged in meaningful instruction are continually testing hypotheses about various relationships, the accuracy of various statements, and alternate ways of accomplishing the task. Educational improvement strategies must include the following four criteria to be successful: (a) Methods must be focused on self-assessment. (b) Educators must have a thorough understanding of value added within each stage of the educational process. (c) Education must broaden the scientific knowledge base for best practices. (d) Schools must employ efforts that empower teachers to test new strategies and contribute to improvement efforts (RAND, 2004). Can TPS expedite this process?

Shuell (1992), describes the components of an effective instructional design process using six components that are based on cognitive learning theory: identify purpose and/or goals, consider the audience/user, specify instructional procedures, assess learner’s knowledge and understanding, provide for alternate instruction, and field test with students to make necessary changes. To effectively identify the purpose
for the lesson, the instructor must identify exactly what the student will be expected to understand or be capable of doing after the instruction is complete. It is important for instructors to know their audience and each member of that audience. “Even students who appear quite similar to the casual observer possess characteristics and exhibit differences that influence the extent to which a particular instructional program is effective (Shuell, 1992, p. 46).” According to Shuell (1992), instructional procedures are broken down into three specific components: presenting the content necessary for the learner to achieve, ensuring that students participate and are motivated, and engaging those psychological processes responsible for learning. The assessment procedure must be considered as part of the instructional process, according to Shuell (1992), and has two specific purposes: information obtained from the assessment can be used to make decisions regarding content and instruction, and the information can provide feedback for the learners to see their growth. Providing for alternate instruction is the final step in this process and is dependent on the assessment information. If a learner is having difficulty with a particular concept, it is necessary to provide additional instruction, but not necessarily in the same manner. “Simply having the learner repeat the same instructional program that he or she has just completed is seldom effective; after all, the first exposure has already proven unsuccessful” (Shuell, 1992, p. 49). TPS offers education a manufacturing-based system that incorporates all six components for effective instruction as described by Shuell (1992).

To effectively utilize the TPS system in education, educators would have to translate the assembly line model into the context of schools. “Although it is not a
perfect analogy, there is a natural overlay of the manufacturing model in the educational context” (Barney & Kirby, 2004, p. 43). In education the workers who directly affect the final product are teachers; instruction and curriculum are jobs located along the assembly line. The plant managers are the principals, superintendents, curriculum director and policymakers, who influence the system. One major difference in the manufacturing-based system and the public school system is the consistency of the starting materials. Business strives to have 100% consistent starting materials, while education strives to produce the same standards from different starting materials. Another difference between education and Toyota’s system is that Toyota employs small, close-knit work teams. Education seems to have teachers in isolation, each trying to deliver instructional units that are consistent with other staff members and standards, which often does not happen (Barney & Kirby, 2004).

TPS utilizes value streams that point out important responsibilities within any given process and focus on the bigger picture, not just inputs or outputs, as is “modus operandi” in a bureaucratic system. Under the TPS model, teachers and administrators would need to address an extended sequence of questions: What final outcomes are desired? What intermediate steps add value that contributes to those outcomes? Where in the production line should each piece of intermediate value be added? Which pieces build on other pieces? What necessary value might currently be missing from the curriculum and instructional process? What steps in the current process do not add value toward the desired outcome (Barney & Kirby, 2004)? It seems that education would benefit by incorporating Jidoka in the testing/learning
process to identify student deficiencies before final tests are given. Utilizing Jidoka in education would put an end to social promotion for those who had not built an adequate foundation in one grade for success in the next (Barney & Kirby, 2004). To create a quality value stream in education, teachers would need to develop valid and reliable indicators.

The teaching/learning process seems equally suitable to benefit from TPS. NCLB calls for a scientifically based practice, which fits perfectly in Toyota’s scientific based design. “In practical terms, TPS suggests that classroom teachers could use standardized instruction and curriculum and, by paying attention to both actual outcomes and expected outcomes, engage in the same sort of hypothesis-testing that Toyota workers engage in” (Barney & Kirby, 2004, p. 47).

Standardization can also be dangerous in education; all students learn in a different manner and at different speeds. This chapter warns against school systems’ implementing too much standardization for reasons of addressing individual student needs. TPS does not, however, endorse mass customization as employed by many other major car manufacturers; standardization is used in TPS to promote continuous learning from experience. It is important that these two concepts are not mistaken as one and the same. TPS also seeks to empower its workers, which should enhance the educational process by training teachers to solve problems and be more reflective about their work.

For the managers in the educational system, TPS suggests that by fostering the advantageous combination of teachers’ specific knowledge with the higher-level expertise and guiding resources provided by researchers, administrators,
policymakers, principals, superintendents, and others could offer support, thus enhancing the process of continuous improvement (Barney & Kirby, 2004, p. 49).

Success in TPS depends upon the dynamic interaction of all three principles, leading to a coherent organizational system in which problems are dealt with at their source, on the lowest level possible, and with continuous and immediate objective feedback (Barney & Kirby, 2004). Much of the literature suggests that for TPS to be successful in creating positive organizational change, it must be implemented in its entirety.

**Measurement Instruments-PSSA**

“In 1999, Pennsylvania adopted academic standards for Reading, Writing, Speaking and Listening, and Mathematics” (Thacker & Dickinson, 2001). Students were assessed at grades five, eight, and eleven in reading and mathematics from 2001-2003. Writing tests were administered during the same years in grades six, nine, and eleven (Thacker & Dickinson, 2001). “The annual Pennsylvania System of School Assessments (PSSA) is a standards-based criterion-referenced assessment, used to measure a student’s attainment of the academic standards, while also determining the degree to which school programs enable students to attain proficiency of the standards (PDE, 2001). Performance level descriptors for the PSSA are labeled *advanced* for superior academic performance, *proficient* to demonstrate satisfactory academic performance, *basic* to demonstrate marginal academic performance, and *below basic* to reflect inadequate academic performance (PDE, 2001). Testing efforts expanded to comply with NCLB for the 2006-2007 school year, and now includes assessment of students in reading and mathematics in grades
three, four, five, six, seven, eight, and eleven. Writing tests were administered to students in grades five, eight, and eleven during the 2006-2007 school year. In accordance with NCLB, science tests are now given in grades four, eight, and eleven. PDE provides school districts with testing windows, during which time the tests must be administered to students. Reading and mathematics tests were given to students in Pennsylvania between March 20 and March 31 of the 2006-2007 school year. The writing test was given to Pennsylvania students between February 13 and February 24 of 2007. PSSA mathematics and reading tests require two and one-half hours to complete. During the 2007 testing cycle, PDE will have identified students scoring 1483 or above as advanced, 1312-1482 as proficient, 1158-1311 as basic, and below 1157 will be grouped as below basic (PDE, Sept. 2006).

The Pennsylvania System of School Assessment (PSSA) test has become a valid tool to measure student achievement in high school and has been used as a predictor for success in further education. PSSA tests are linked with many different variables in school systems. “Pennsylvania school library programs can make a difference supporting the efforts of schools to measure up to standards” (Lance, Rodney & Hamilton-Pennell, 2000, p. 6). Reading scores increase when school districts increase funding in areas of the library in regard to staffing, technology and incorporating information literacy into the school curriculum. PSSA reading scores are especially connected to the number of computers made available to both teachers and students who have access to the Internet. “Test scores increase as school librarians spend more time: teaching cooperatively with teachers; teaching information literacy skills independently; providing in-service training to teachers;
serving on standards committees; serving on curriculum committees; and managing
information technology” (Lance, Rodney & Hamilton-Pennell, 2000, p. 7.) With
these specific school library programs improvements, PSSA reading scores tend to
increase, on average, 10 to 15 points.

PSSA test results have also been a predictor of college success and various
alternative assessments. The Human Resources Research Organization (HumRRo)
investigated whether or not there is a relationship between PSSA test scores and the
SAT tests or other local assessments (Shorr, 2005). Shorr (2005) reported findings
that indicated a high correlation between how a student performs on the eleventh-
grade PSSA test and the SAT test. By using predictive validity, the PSSA test can
help to predict how a student will succeed in his or her freshman year of college. The
findings encourage students to use the results of the PSSA tests to help address their
needs during their high school career, before they enter post-secondary education.

Educational accountability has been a focus in the development of
assessments such as the PSSA. These high-stakes tests are deemed a reliable measure
of what students are being taught. Results of these tests are defined as bottom line;
they define the success or failure of the student and the school (Zwerling, 2001).
Various events can occur if a score of nonproficient is obtained. Students who do not
pass the PSSA test may undergo content remediation, changes in class placement and
in course flexibility. Schools that do not meet AYP may lose employees and be
forced to implement new curricula. The situation could also lead to privatization.

The Philadelphia School District put into practice a mathematics intervention
that targeted students in grades three through eight. The In Math program was
The Toyota Way

designed to: strengthen basic skill, contain a technology element, and create competition to motivate students to achieve (First In Math, 2004). During the first year of the program, the students logged in over 216,000 hours and increased their PSSA tests results 7.4% in fifth grade and 11.1% in eighth-grade.

Zwerling (2001) performed a study that analyzed the relationship between PSSA test performance levels and that of other commercial assessment tests. Although there proved to be a strong correlation between PSSA test scores and other assessments, many students who scored low on the PSSA tests, scored higher on other tests such as the SAT’s.

The PSSA test and other various assessments have been used to facilitate increases in student achievement. Governor Rendell stated that Pennsylvania is one of only seven states that made improvements concurrently in both reading and math programs in the elementary level (Pa PowerPort, 2006). From 2002 through 2005, more than 8,126 additional fifth-graders were proficient in reading and 19,938 more were proficient in math. “The Governor’s vision is clear: every student will achieve proficiency – no exceptions, no excuses” (Pa PowerPort, 2006, ¶ 9). Bedford Elementary School has been recognized by Standard & Poor’s as one of the districts that reduced the achievement gap over the past three years.

4Sight Benchmark Assessments

4Sight Predictive Benchmark Assessments were designed by the Success for All Foundation and the Center for Data-Driven Reform in Education at Johns Hopkins University (Chute, 2006). Testing options for the 2006-2007 school year include reading and mathematics tests in grades three through eleven. 4Sight
Predictive Benchmark Assessments are scored on a 30-point scale that can be translated into PSSA categories of advanced, proficient, basic and below basic. The tests are designed to be low-stakes, formative assessment tools, used to guide decision making and instructional reform efforts (Success for All, 2006). 4Sight assessments are one-hour tests that have exactly the same formats, coverage, look, and feel as state assessments (Success for All, 2006). 4Sight is a standards-based criterion-referenced test that models questions, percentage of question type, and formats used on the PSSA test (Success for All, 2006). With NCLB holding districts accountable for every student to perform at a higher standard, and for all students to achieve proficient performance, school districts are turning to data for answers. One of the most important issues for educators, as stated earlier in this literature review, is obtaining reliable data in a reasonable time span. 4Sight tests were designed to provide schools with reliable data and immediate feedback. These data would be reliable in predicting student performance on the PSSA, and would also provide teachers with formative type information for instruction (Chute, 2006). These assessments are one hour in length; standards-based, formatted in the same style as the PSSA, and have equivalent percentages of question types. They address standards similar to the state assessments (Success for All, 2006). Schools cannot afford to wait for results of the PSSA to see how students will perform in their elementary and middle schools. “That is why the Success for All Foundation created 4Sight, a new benchmark assessment tool that enables you to predict your students’ achievement five times a year” (Success for All, 2006, ¶ 1).
The Pennsylvania Department of Education (PDE) presented “Getting Results,” a revised framework for school districts to follow for continuous school improvement. In this document, PDE stresses the need for districts to utilize data from multiple perspectives to “paint a portrait” of the current state of student learning (PDE, 2006). This 45-page document details action steps for school districts to follow, including a section entitled “Key Components.” Within this section, PDE lists 4Sight assessments as one of three formative assessments that school districts should be using to predict student performance and guide instruction. The “Getting Results” document also includes a worksheet for districts to analyze 4Sight results according to each state standard (PDE, 2006). 4Sight tests and similar measurements are termed “benchmark assessments” (Olson, 2005). These benchmark tests are administered throughout the year to measure student and district progress in preparation for the PSSA tests in the spring. Benchmark assessments are aligned to state or district standards to help teachers better plan and adjust teaching instruction and curriculum.

High-stakes testing has become so prevalent that the director for Data-Driven Reform in Education, at Johns Hopkins University, stated that in three to five years, people will not remember a time when there were not benchmarks (Olson, 2005). Companies that produce these benchmark tests hope that the trend of testing continues. It is predicted that in the year 2006, nearly districts for such tests would spend 320 million dollars.

Some of the concerns pertaining to the 4Sight tests are the tests ability to be a predictor of a student’s level of achievement at the end of the school year. 4Sight tests do not predict end-of-year achievement; however they do measure actual progress
that students are making from test to test (Claycomb, 2006). These tests are designed as quarterly benchmark tools to provide useful data for focusing professional development and instructional goals (CDDRE, 2006).

The validity of 4Sight Benchmarks was analyzed on the basis of its predictability on student performance on high-stakes assessments. “The 4Sight Reading Benchmarks for Pennsylvania, 2004-05, were correlated with grades three and five of the PSSA. The correlations ranged from .83 to .89” (4Sight Technical ReportPA, 2006, p. 23). The correlations are an indicator that the 4Sight benchmark assessment tests are a good predictor of student achievement on PSSA tests.

Ohio State Teacher Efficacy Scale

Teacher efficacy is paramount in helping a school reach the highest achievement possible. It has been used to describe teacher quality, as well as to gauge teacher change. Teacher efficacy demonstrates a person’s perception and confidence in their capacity to encourage student learning (Shore, 2004). Teacher efficacy seems to be directly related to student achievement. Teachers perceive themselves as more successful and competent as their students experience increased levels of achievement.

According to Shore (2004), teacher efficacy has been found to be associated with many powerful forces in teaching and learning, including, but not limited to, the following:

- A sense of personal accomplishment, where teachers view their work as important
- A willingness to try innovative practices
• Personal responsibility for student learning in that area
• Strategies for achieving objectives for their students
• More persistence with students who struggle or have special needs
• Greater job satisfaction, which correlates with greater retention
• A sense of control in the classroom or a belief that the teacher can influence student learning
• A sense of common teacher/student goals and democratic decision-making

“Examining a teacher’s efficacy as a result of a program or intervention is one way to evaluate the effectiveness of the program” (Shore, 2004, p. 116).

Yost (2002) completed a study to determine whether mentoring may be used as a means to enhance teacher efficacy. “To effect classroom change, teacher characteristics that most directly affect student learning, including communication skills, instructional style, planning and management skills, and content knowledge must be examined” (Yost, 2004, p. 195). When learning teachers understand styles of students, instruction may be designed appropriately; the same is true for teachers and administration. Administrators must better understand what teachers’ needs are in order to enhance educational reform.

In Yost’s study (2002), mentor teachers where chosen on the basis of first-rate teaching performances. Teachers took a mentoring class the summer previous to the start of the school year, preparing methods to enrich their fellow teachers. The mentors were given an entire year off from teaching to provide ample time for enhancing the mentees’ abilities. Reflections from the study demonstrated growth by all individuals involved in the mentor program.
The mentors who participated in the program stated that they became more aware of their teaching and of the responsibilities they had to their students. Looking at classroom learning through the eyes of another often resulted in new realizations about how teacher practice could directly affect learning (Yost, 2002, p. 196).

Both the inexperienced and master teacher learned from the opportunity. Shared learning opportunities allow school districts to advance in their efforts to employ highly qualified staff in every classroom.

Ebmeier (2003) states “Although formative teacher evaluation, often called supervision, is a common feature of life in schools, very little is known about its effect on teachers or the mechanism by which supervision affects instruction”(p.110). Ebmeier (2003) designed a study that tested how or if supervision affects teacher efficacy. Results of the study indicate that supervision has a profound impact on teachers’ levels of commitment to teaching and teacher efficacy. The part that seems to play the biggest role is a person’s trust in his or her peers, as well as in the building’s administration. A principal who is out and about in the hallways does not specifically enhance teacher efficacy. Principals need not only to be visible, but also be viewed as trustworthy and attached to teacher/student activities. When the administration is seen as taking on the projects of their building to heart, teacher efficacy seems to follow in a positive manner. According to Ebmeier (2003), “If teachers believe they can overcome external conditions, this will strongly influence beliefs about their own teaching competence” (p. 140).
Culture is another item that is examined closely in regard to teacher efficacy. “Teachers assume a critical role in creating classroom environments that encourage students to become active, self-motivated learners” (Deemer, 2004, p. 73). Teachers and students alike need to feel that the culture of the school is supportive to the instructional practices used in the classroom. School culture should be in compliance with high educational standards and the classroom should present to the same type of atmosphere. “Teachers’ personal beliefs about their capabilities to help students learn, or personal teaching efficacy, have been shown to influence the goals teachers promote in their classrooms” (Deemer, 2004, p. 74). Teachers and students teach and learn to the abilities that are expected of them. If teachers demonstrate low levels of competency in their teaching ability, they often fail to challenge their students with higher-level activities. Teachers with high-levels of efficacy are more likely to challenge their students with educational materials, as well as encourage student understanding to those having difficulties.

Teacher efficacy is determined by using a Teacher Efficacy Scale (TES) that was first designed by Bandura and revised by Woolfolk-Hoy (Shaughnessy, 2004). It was stated in the interview with Woolfolk-Hoy that Bandura felt the term “teacher efficacy” was confused with effectiveness. Bandura preferred the terms “teachers’ sense of efficacy, self-efficacy of teachers, instructional efficacy, teachers’ efficacy beliefs, or teachers’ perceived efficacy” (Shaughnessy, 2004, p. 154). Some of Woolfolk-Hoy’s findings indicate that each teacher is affected differently in regard to personal efficacy. Teachers’ perceptions of their own efficacy are dependent upon what subjects they are teaching, the size of their classes, and their knowledge of their
students (Shaughnessy, 2004). In an effort to enhance the existing Bandura’s efficacy scale, Woolfolk-Hoy added items that would capture the important tasks of teaching and eliminate trivial information.

There are four sources that Bandura identified for efficacy beliefs: modeling, mastery experiences, persuasion, and physiological arousal. Woolfolk-Hoy added teacher self-regulation strategies to Bandura’s beliefs. “Since teacher efficacy is a critical variable in studying many educational concerns, it is of great importance to adapt the TES so that it becomes a valid measurement instrument” (Brouwers & Tomic, 2003, p. 78).

The Ohio State Efficacy Scale is the instrument chosen for this research study. Anita Woolfolk-Hoy of Ohio State University and Megan Tschannen-Moran of William and Mary College developed the tool. According to Megan Tschannen-Moran, there are fluctuations in teacher self-efficacy any time they are exposed to new ideas (Moran, 2006). The instrument is being used more often now and has proven to be most valid when used with in-service teachers because of the option of three scales (Moran, 2006).

Summary

With the growing demands placed upon schools by NCLB, schools are forced to utilize data to make important decisions regarding student performance, teacher performance, and instruction. Ignoring the facts that schools will be held accountable for student performance and that schools will be responsible to show student improvement, is a fatal mistake. Educators must find ways to utilize data that are relevant, readily available, and applicable to decision-making.
To effectively utilize data sources for decision-making, schools must first build a professional culture where staff members understand testing results and data. Building such a culture is a pillar around which TPS has built its system. Effective use of data transpires in the context of a robust professional learning community, where teachers and administrators are crystal clear about their vision and their commitments, relentlessly focused on results for students, collaborative and reflective about their practice (Love, 2004). Collaborative structures must be created to allow teachers time to analyze data resources and apply them to instruction. If teachers are going to use data, generate strategies to improve student learning and monitor the results, they need time to meet weekly in department meetings, vertical teams, grade-level teams, or study groups (Love, 2004). Under TPS, the teachers would receive the data in real time and as needed.

Teachers should be provided opportunities to engage in data-influenced dialogue so they can make collective sense of the data and take ownership in the problems and the solutions. Standardized tests provide important information that can be used in the decision-making process, but it must be clearly understood that standardized tests results are not the only type of data needed for improvement. Teachers should also use multiple measures, including common grade-level, subject area, or course-specific assessments. “One key to the district’s success is the use of common assessments designed by teachers to assess the knowledge and skills the teachers agreed were central to their curriculum” (Love, 2004, ¶ 2). “It is the data-influenced dialogue that takes place in department, course, or grade-level teams, not
the rank-ordering of schools in the newspaper, that provide the real momentum for improving student learning” (Love, 2004, ¶ 4).

It becomes increasingly clear that educators must utilize data derived from relevant sources that provide accurate information for making decisions. Teachers and staff members need adequate time to analyze the data and make necessary adjustments in the instructional process. Standardized tests provide teachers and parents with information that is helpful in the process of improving schools and measuring student success. It is equally important that schools learn processes that work in order to achieve management goals that have been derived from data, and continue to grow in a culture of continuous improvement.
CHAPTER THREE

METHOD

Introduction Statement

The Bedford Area School District entered into a grant agreement with The Grable Foundation to implement TPS into two fifth-grade math classes during the 2006-2007 school year. Bedford included into the grant application the necessary funds to have a TPS consultant aid the district in the implementation process. The primary purpose of this retrospective case study was to determine if the management tools used in the TPS model improved student achievement in two fifth-grade elementary math classrooms and created changes in teachers’ perceptions of self-efficacy for those fifth-grade teachers trained in TPS principles. In addition to the major research agenda regarding achievement and efficacy, this research describes the implementation using the experiences of the researcher, classroom teacher and the TPS implementation consultant, Christina Dixon.

Dixon is the educational consultant that aided the Bedford School District’s implementation process for TPS. Dixon was a partial owner of True North Consulting prior to the company’s closing in the summer of 2006. Even though the company closed prior to the start of this project, Dixon decided to stay with the Bedford project, seeing it through to completion. True North had provided limited business consulting for carefully selected clients. True North was also a TPS University, providing a unique and in-depth introduction to TPS principles for organizational leaders (True North, 2005). True North Institute, a service offered by True North,
The Toyota Way

provided leadership to the TPS University for some of its clients. TPS University provides five-day training sessions for leaders of organizations adopting TPS. The training was closely modeled after the executive training that Alcoa developed for all its business unit presidents when it adopted TPS company-wide. At True North, the training was specifically tailored to meet specific client needs in various sectors. True North Institute had partnered with agencies to implement models of TPS into healthcare and other nonprofit sectors. True North has implemented TPS principles in industry, healthcare and education. They have applied TPS principles at Alcoa Corporation, Giant Eagle, the Women’s Center and Shelter of Greater Pittsburgh and Pittsburgh area hospitals (True North, 2005).

As one of three founders of former True North, Dixon led the Bedford School District’s implementation of TPS management and data-influenced decision-making in two selected fifth-grade math classrooms. The purpose was to determine if the principles used in TPS significantly increased student performance as it related to PSSA state testing and 4Sight predictive benchmark assessments in mathematics. The researcher also examined existing data to analyze any changes in teacher efficacy, as indicated by OSTES, which may have occurred as a result of the TPS implementation. Fourth- and fifth-grade teachers completed the OSTES as part of the Grable grant, prior to the start of the 2006-2007 school year. The researcher in this study analyzed existing data from the same group of teachers, using the OSTES, at the end of the 2006-2007 school year. This information was used to determine if the fifth-grade teachers who have been trained or exposed to TPS principles, showed
significant gains in teacher efficacy, when compared to fourth-grade teachers who had not received instruction in TPS principles, as indicated by OSTES.

**Target Population**

The target population was fifth-grade public school students from Bedford Elementary, a school located in a rural district in Bedford County, Pennsylvania. The Bedford Area School District has a total enrollment of 2,319 students in grades kindergarten through twelve. The target populations of fifth-grade students are located in the main elementary facility that houses approximately 810 students in grades kindergarten through fifth. The focus was on the entire fifth-grade class of 141 students, during the 2006-2007 school year, which completed the PSSA exam during the previous 2005-2006 school year. Student test scores from the PSSA state tests and 4Sight benchmark assessment tests were used for the dependent variables. Particular attention was focused on two TPS math classes, totaling 45 students, who were instructed using data-influenced decision-making through the TPS design, guided by Dixon. Existing teacher data from the OSTES, collected at the start of the Grable grant, in August of 2006, and post grant surveys collected by the district in June 2007, were used to create a comparison of changes in teacher efficacy brought about by TPS training in the fifth grade.

**Method of Sampling**

The sample was a cohort of students who were fourth-graders during the 2005-2006 school year and fifth-graders during the 2006-2007 school year that took the PSSA test during both years. Fifth-grade students from the Bedford Elementary School were the sample population. All 141 students in the fifth-grade took the 4Sight
progress monitoring tests given five-times throughout the 2006-2007 school year.

Focus was placed on measuring performance changes on the PSSA for all students in grade five who took the PSSA test in grade four and on all fifth-grade students using scaled scores from 4Sight Predictive Benchmark Assessments. Particular attention was directed toward the TPS mathematics classes to observe how TPS impacts student performance, as indicated by PSSA and 4Sight Predictive Benchmark Assessments. The researcher used existing data from the OSTES, collected at the start of the Grable grant in August of 2006 from fourth and fifth-grade teachers and from post study OSTES survey information from the same groups of teachers, to analyze teacher efficacy changes. Teacher efficacy demonstrates a person’s perception and confidence in his or her capacity to encourage student learning (Shore, 2004).

Teacher efficacy can demonstrate if teachers feel their work is important and show their willingness to try new practices. According to Shore (2004), examining teacher efficacy, as a result of program or intervention, is one way to evaluate the effectiveness of the program.

**Stimulus Materials**

TPS management strategies were implemented into a classroom model, designed to instruct 45 students, in two separate fifth-grade mathematics classrooms, at the Bedford Elementary School. The same teacher who was instructing students using data-influenced decision-making through the TPS management system taught both classes in the experimental group. The TPS management system classroom design and instructional process was the single targeted stimulus. The effects of TPS on student achievement was measured using the 2006-2007 fifth-grade PSSA tests.
from the same cohort of students that took the PSSA in grade four during the 2005-2006 school year and 4Sight Predictive Benchmark Assessment scaled scores from all students in the fifth grade during the 2006-2007 school year. Since students were not selected from any given academic program, 4Sight tests results were analyzed for the entire fifth-grade population at Bedford Elementary.

TPS influenced the targeted classroom teacher’s instructional decision-making and classroom management processes as suggested by the TPS implementation consultant and as mapped in the Classroom Design Model, Appendix C. The design of the model classroom was set up to create a process flow that surfaced problems by identifying any students who do not reach the targeted achievement level of 85%. Checkpoints were set up throughout the instructional process, allowing the teacher to monitor student results using any of the following tools: observation, math facts quiz results, teacher designed tests, chapter tests and 4Sight Predictive Benchmark Assessments. Students who were identified as achieving at a lower level than 85% were prescribed maintenance and/or remediation. Remediation consisted of reteaching the material to a particular group of students, prescribing them to specific tutoring, utilizing computer programs to reinforce math facts and/or repeated practice sessions for identified skill deficiencies. If these remediation efforts did not result in the student achieving at the 85% level for a particular skill, the teacher then looked for the root cause of the learning problem, which may have required the use of the Help Line Flow Chart, Appendix D. This process requires the teacher to see the big picture of a child’s learning process and creates teacher accountability for each student achieving at the 85% level on all assessments. Within the TPS model, a
teacher must work toward continuous improvement for all students to achieve at the targeted 85% level on all assessments, and the teacher must maintain contact with the help line to assure that problems are being addressed in a timely manner. It is not acceptable for any student to achieve at a level of less than 85% proficient. The TPS model classroom design is set up in such a way that the culture of the learning environment is changed, with the primary focus being on student achievement and the classroom teacher being more accountable for student results and problem solving.

Classroom changes included more upfront preparation during the summer and prior to the start of school. During the summer of 2006, three fifth-grade teachers spent a total of 13 days preparing for the 2006-2007 school year and the TPS model mathematics classroom. TPS requires remediation for any students who are not achieving at the 85% level, and it also prescribes enrichment for students who are at or above the prescribed level of 85%. Enrichment and remediation materials must be readily available as needed by the student, or customer, during the educational process. Teachers spent 13 days designing materials using research-based strategies and tools that were available to the district. Teachers designed math facts practice kits for addition, subtraction, multiplication and division. Individual student tracking folders had previously been prepared for each student in the TPS mathematics classes. Various learning games were assembled into usable units for both remediation and enrichment. An assortment of games and learning tools used in the TPS classroom included, but were not limited to: Solve the Math Mystery, Moving with Tangrams, Eight Versa Tiles activity books, Leveled Problem Solving pages,

Summer workdays are listed in Appendix A, along with specific work times and an accomplishment list. Student PSSA test scores were reviewed prior to the start of school to determine customer need and to obtain relevant data for decision-making. Student learning packets were prepared prior to the start of the 2006-2007 school year to allow for individual enrichment or remediation instruction to take place at any given time, and on demand, as assigned by the teacher. Teachers involved in the preparation also developed a student math-learning center that was available to all fifth-grade students for individual remediation, which took place as necessary for students not meeting the targeted goals. As part of the adaptations to TPS, a classroom store was developed during the 2006-2007 school year by the same three teachers, under the guidance of Dixon, to provide meaningful research-based learning opportunities for any of the following situations: remediation, enrichment, tutoring, practice and evaluation. The model line implementation plan may be viewed in Appendix B. The entire classroom design is based on the scientific approach to problem solving, as used in TPS. This classroom design model can be viewed in Appendix C.

The consultant and the fifth-grade teachers also established a help line for the experimental groups’ classroom teacher to use in the event that students were not reaching its targeted goals. TPS utilizes a similar help line on its assembly line. Administrators and selected teachers were included in the help line flowchart attached in Appendix D. TPS has a standard procedure to identify a problem that would
require the use of the established help line included in Appendix E. Appendix F is a problem documentation log that identified the date of the problem, what person identified the problem, a description of the problem and the date the problem was addressed or corrected. The purposes of the help line and the problem documentation log are to gain a common understanding of the work done to date, examine the goals of the model classroom and adhere to the model classroom design, as indicated in Appendix C. The help line established safety for the classroom teacher and ensured timely response times from the individuals listed on the flow chart. Select administrators and teachers were included on the help line for use when immediate assistance was needed in order for the classroom teacher to meet targeted student goals.

*Measurement Devices- PSSA*

One of two measurement devices that were used for determining student academic progress was the PSSA scaled score test results issued by the Pennsylvania Department of Education. NCLB is a national effort to improve school performance and create more accountability for school systems. As a result of NCLB, states were required to create a system of testing that would accurately measure student performance as indicated by academic standards. PSSA tests are aligned to the Pennsylvania academic standards, which represent expected performance levels for all students and schools in Pennsylvania. Students are assessed in reading and mathematics in grades three, four, five, six, seven, eight and eleven. Student results are reported to parents and schools according to performance levels of: advanced, proficient, basic and below basic. This standardized test is given to all students in
Pennsylvania and provides reliable measure of student performance as it relates to the standards for the commonwealth of Pennsylvania (Koger, Thacker & Dickinson, 2004).

The Human Resources Research Organization of Alexandria, Virginia, was hired to conduct a study to determine the reliability of the PSSA test scores in grades five, eight and eleven. This research was instigated by criticisms directed toward the standardized testing movement in Pennsylvania. The data concluded that the test does not measure all standards, but does measure a good cross section of the standards without focusing too heavily on any one standard. Students’ PSSA test scores strongly matched their performance on other commonly used standardized tests, such as the Terra Nova, the California Achievement Tests, and the SAT. There is a particularly high correlation for the SAT math test. PSSA achievement gaps based on race, gender and economic disadvantage were similar to those on other standardized tests (Koger, Thacker & Dickinson, 2004).

Maher (2003) completed a study at The Center for Education at Widener University that showed similar correlations between the PSSA test results and other standardized test results. The study was conducted in February of 2003 and was titled, The Predicted Validity of the Pennsylvania School System of School Assessment Using the CTB McGraw-Hill Terra Nova Test. The purpose of the study was to determine if the results of the seventh-grade Terra Nova tests could be used to predict the scores of the eighth-grade PSSA scores for the same students. The study utilized samples of student scores from three consecutive years on both the PSSA and the Terra Nova tests. The results indicated that there was a predictive correlation between
the seventh-grade Terra-Nova math and reading assessments, and the PSSA math and reading assessments (Maher, 2003). Results also indicated that males scored better than females on both the Terra-Nova math and the PSSA math (Maher, 2003).

**4Sight Benchmark Assessments**

The second dependent variable was the 4Sight predictive benchmark assessments for math. The 4Sight benchmark tests were developed by the Success For All Foundation and the Center for Data-Driven Reform in Education at Johns Hopkins University to provide a predictive measure of each student’s performance on statewide assessments, specific to the individual state (Chute, 2006). The 4Sight assessments are designed to be formative in nature, while providing schools with predictability data for PSSA tests. The benchmark assessments are given to students up to five times per year. The 4Sight tests assessments are formatted exactly the same as the PSSA tests; they cover the same topic material, and they look and feel like the PSSA exam (Success For All, 2006). Each of the five tests takes approximately one hour for students to complete and was formatted in A/B test design style. The A/B testing format allows for multiple tests to examine the same standards through a variation of question types, therefore creating a measure of a student’s ability level specific to identified skills and standards. 4Sight benchmark assessments are listed by the Department of Education as one of the “key components” to continuous improvement planning for schools (PDE, 2006).

The two types of data provided by 4Sight benchmark tests include student proficiency levels and subscale data. Benchmark assessments are designed specifically to measure the same standards as PSSA and to model the design of each
type of test question. Appendix H provides a comparison of Bedford’s actual student results from 4Sight benchmark assessments and the PSSA test results for grades three, four and five from the 2005-2006 school year. These data were collected and charted by the Bedford Area School District during the summer of 2006. As shown in Appendix H, 4Sight predictions are very closely correlated to actual test results, respectively, across three different grade-levels and in two different buildings.

During the 2005-2006 school year a total of 233 districts across the state utilized 4Sight assessments in grades three through eight; during the 2006-2007 school year, that number expanded to 303 districts and covered grades three through eleven (Chute, 2006). According to a release by PaTTAN (Pennsylvania Training & Technical Assistance Network) in October of 2006, “The Pennsylvania 4Sight Benchmark Assessments are valid, reliable and aligned to the PSSA and provide an accurate measure of student performance on the PSSA, as well as diagnostic sub-skill data to guide classroom instruction and professional development efforts” (PaTTAN, 2006, p. 1). Since this research study used the 4Sight tests to measure the mean changes in student scores between the experimental and the control group, predictability is not a concern.

Ohio State-Teacher Efficacy Scale

This study shows changes in teacher efficacy resulting from the experimental groups’ training using TPS principles. Prior to the start of the Grable grant, all teachers at Bedford Elementary in grades four and five completed a teacher efficacy instrument that was designed in 2001 by Megan Tschannen-Moran from the College of William and Mary and Anita Woolfolk-Hoy from Ohio State University. A copy of
the instrument is included in Appendix G. This instrument was developed at the Ohio State University and will be referred to as the Ohio State Teacher Efficacy Scale (OSTES).

In 2002, the OSTES was used in a Tennessee State University study to determine the relationship between high-stakes testing, teacher efficacy and school performance. The efficacy scale was given to 83 educators from three low-performing high schools and 48 educators from high-achieving schools. The factor analysis revealed three factors: instructional strategies, classroom management, and student engagement. A MANOVA revealed through the Wilks’ Lambda multivariate test that there was a statistical significant difference between the performance level of schools and the three factors associated with teacher efficacy, as well as a relationship between gender and school performance level. “The Ohio State Teacher Efficacy Scale proved to be both reliable and valid based on the results from the factor analysis and Cronbach’s Alpha” (Smith, 2002, p. iv).

Construct validity of the OSTES was tested in three separate studies to confirm the reliability for both the short and long forms by evaluating the correlation of this scale in comparison to other existing measures of teacher efficacy (Tschannen-Moran & Woolfolk-Hoy, 2001). In a principal-axis factor analysis, used to determine the appropriateness of calculating a total score, the reliability of the 24-item scale was .94 and the 12-item scale was .90. The results for both the long and short forms were positively related to the personal teaching efficacy and the general teacher efficacy. The strongest correlation was between the OSTES and personal teaching efficacy. Less of a correlation existed between measuring general teacher efficacy and the
essence of efficacy. Total scores on the OSTES were positively related to both the RAND items (r = 0.18 and 0.53, p< 0.01) as well as to both the personal teaching efficacy (PTE) factor of the Gibson and Dembo measure (r = .064, p< 0.01) and the general teacher efficacy (GTE) factor (r = 0.16, p<0.01) (Tschannen-Moran & Woolfolk-Hoy, 2001). “The results of these analyses indicate that the OSTES could be considered reasonably valid and reliable . . . Positive correlations with other measures of personal teaching efficacy provide evidence for construct validity” (Tschannen-Moran & Woolfolk-Hoy, 2001, p. 801).

The studies revealed weaknesses in other efficacy scales, which led Tschannen-Moran and Woolfolk-Hoy to develop a more supportive assessment that addressed a broader array of teaching tasks. The three dimensions of efficacy that were added to the OSTES scale were instructional strategies, student engagement, and classroom management (Tschannen-Moran & Woolfolk-Hoy, 2001). As stated earlier in Chapter Two, Moran (2006), confirms that this scale is better used with in-service teachers because the three scales listed above are valid measures. Moran (2006) also confirms that this scale is less valid when used with pre service teachers, or teachers who are not currently active. For the Bedford TPS implementation, the OSTES measured efficacy for full-time teachers. The most obvious limitation to the efficacy results was the small sample size of the experimental group.

Data Collection Methods

This study utilized a computer-based analysis system to collect data for both the PSSA results and the 4Sight benchmark assessment tests. The Bedford Area School District’s computer data analysis staff, using the eMetric or CDA, collected actual
student data. Student names were removed from the data with the only identifying factors being their grade level and if they were in the TPS model classroom. Students were assigned a number to match their fourth-grade PSSA test results to the same student’s fifth-grade test results. PSSA results are reported from the state to the districts late in the summer of each school year. The Intermediate Unit then places the test scores into a database called Computer Data Analysis (CDA). This database allows school systems to make comparisons of various data, not limited to but including PSSA test scores. Data from the 4Sight progress monitoring tests were obtained from the 4Sight homepage online reporting program available to the districts that purchase 4Sight-testing instruments. Student names were removed from the data source by the Bedford Area School District’s data analysis staff. Students were identified by grade-level, with a numerical value to match test results. This study also examined teacher efficacy instruments that the fourth and fifth-grade teaching staff completed in August of 2006, and then again in May of 2007. Teachers are identified only by the grade-level for which they teach and if they were in the TPS implementation program. Teachers were not identified by name; the only identifying factor was their grade-level of instruction and whether or not they taught in the TPS model classroom. Once the data were collected for both PSSA and 4Sight, they were entered into SPSS for final analysis and reporting.

Research Design and Procedures

The TPS model classroom design for instruction was one stimulus for change in student test scores. TPS model design makes teachers accountable for student achievement and for problem solving to the root cause. The two dependent variables
were PSSA scaled state test scores and 4Sight Predictive Benchmark Assessment scaled scores. The second stimulus was teacher training in TPS principles and classroom management model design. Teacher efficacy was measured using the OSTES to determine if the fifth-grade teachers showed any statistical significant group mean score changes in their perceptions of teacher efficacy when compared to fourth-grade teachers who did not receive TPS training. Inferential statistics could not be used due to the small sample size and inability to make conclusions about the data. The third dependent variable, a teacher’s efficacy scale (OSTES), was derived from existing data collected prior to the start of the Grable grant in August of 2006 and from a survey given by the researcher in June of 2007.

The repeated measures one between/one within design used for comparing student 4Sight scores allows for each individual in both treatment conditions to be analyzed, which, in this case, are the experimental group receiving instruction through TPS and the control group. This study analyzed data from August of 2006 and ended in June of 2007. “The repeated measures one between/one within design is especially well suited for studying learning, development, or other changes that take place over time” (Gravetter & Wallnau, 2004, p. 355). Using repeated measures design allows for individuals to be measured at one time and later the same individuals measured again, allowing the observer to examine what changes took place during the elapsed time. “The primary advantage of repeated-measures design is that it reduces or eliminates problems caused by individual differences such as: age, IQ, gender, and personality” (Gravetter & Wallnau, 2004, p. 356).
Repeated measures F-ratio was calculated using SPSS to determine if the independent variable (TPS) had a significant effect on the dependent variables listed previously in this section. The F-ratio was calculated by measuring the treatment effect, adding to it the error and dividing the sum by the error. By using the repeated measures design, variability due to individual differences in participants was eliminated, thus allowing for a more accurate F-score. “In statistical terms, a repeated-measures test has more power than an independent-measures test; that is, it is more likely to detect a real treatment effect” (Gravetter & Wallnau, 2004, p. 463).

In a first analysis, the PSSA scores between the experimental group and the control group will be compared using 2005/2006 and 2006/2007 student results. A two way mixed analysis of variance with a between-subjects factor of group and a within-subjects factor of year will be used to analyze the data. A second analysis of 4Sight student data results will use a repeated measures analysis of variance with a within-subjects factor of occasion and a between-subjects factor of group to test the change in math scores. Changes in mean scores between occasions will be tested using paired samples t-tests. The researcher will also utilize existing teacher efficacy data, collected by the Bedford School District, from fourth and fifth-grade teachers in August of 2006 to compare with data that will be collected from the same teachers in June of 2007. The design for this analysis will be a two by two mixed anova with two independent variables being TPS training and pre-post scores. In addition, this study sought to clarify findings by telling the implementation story with a descriptive analysis of the study through the eyes of the researcher, the implementing teacher and the TPS consultant.
TPS Implementation Process

TPS was first discussed in the Bedford School District on September 26, 2005, at which time the district superintendent and assistant superintendent set up a meeting of possible key persons from each building in the district. This initial meeting was conducted with sixteen staff members including the superintendent, assistant superintendent, seven elementary teachers, three secondary teachers and four district building level administrators. The purpose of the meeting was to discuss the possibilities of implementing an industrial based management system into an educational setting at one of the Bedford schools. Discussion revolved around possible funding sources for the project, identifying the key members for the implementation, focusing the team on a specific problem, and most important, to identify a possible goal for the project. During this first meeting, the TPS values and principles were discussed through case study analysis and simulation activities. The first meeting was designed to generate thinking and to allow team members to decide whether or not to be part of the process.

On Monday, October 17th of 2005, the team met for a second time to learn more about what TPS is and what other organizations have done with the TPS process. At this meeting the team of Bedford educators met with members of Duquesne University and with Dixon. Duquesne University representatives and Dixon unified to help initiate the innovative concept of integrating TPS into an educational setting. During this second meeting, the team reviewed the key concepts of TPS, as described in Chapter Two, and searched for possible insights on improving the educational process at one of our schools. Considerable effort was focused on
determining the present condition of our district. Team members observed video clips of teaching and learning and then divided into focus groups to discuss key elements of the process. Large group discussion revolved around setting a focus for the implementation and identifying a teacher to accept the challenge. The afternoon session involved watching video clips from the Women’s Center and Shelter of Greater Pittsburgh, a service organization that implemented TPS with success. The executive director of the women’s center was also present to explain the process that their organization experienced and how it helped to make their organization more effective. At the conclusion of the second meeting, the Bedford team decided to continue with the TPS effort by having Dixon conduct observations and interviews in the district from November through March, during the 2005-2006 school year. The purpose of Dixon’s work was to identify improvement opportunities in the Bedford School District that may serve as a target goal for TPS. The team agreed to reconvene prior to the start of the 2006-2007 school year to discuss what a first year pilot may look like. While the initial implementation work was taking place at Bedford, a group of doctoral students from Duquesne University was working to put together a proposal to find potential funding of the project. This completed document was given over to Duquesne University faculty and the Bedford School District Superintendent in the spring of 2006.

Determining the current condition consisted of a series of observations and interviews of over 60 staff members, parents, students and board members from the Bedford Area School District. Dixon spent November through March’s 2005-2006 school year traveling around the district to learn how the business of educating
children is conducted in Bedford. Her initial activities consisted of examining the indicators of success to gain a better understanding of the district’s overall vision for educating children. By observing teachers in action, Dixon was able to gain a better understanding of the system from the teacher’s point of view and where the actual education is taking place. Throughout this observation process, Dixon sought to understand how teachers are connected to other staff members within the system and how they work together to educate children. The final steps needed to finalize Dixon’s evaluation were to analyze the data that had been collected and examine the system using the TPS structure as a guide. Dixon prepared a recommendation for the first year pilot project that would designate the elementary math department as a target area for the 2006-2007 school year.

On April 10, 2006, the Bedford team reconvened to review the case study report and to establish a target condition using a comparison chart developed by Dixon. This meeting marked the third meeting of the district’s team. The focus of this meeting was to create a vision for the implementation process during the 2006-2007 school year. Dixon, along with faculty from Duquesne University, led the team through the findings produced from the interviews and observations that had been conducted during the previous school year, which established the present condition. Once the present condition of the district had been established, it was time to focus the group on the target condition and the implementation process. It was the desire of the team to focus on student learning as it related to students scoring proficient or advanced on the PSSA exams. Dixon led the team in selecting a small slice of the problem, which eventually indicated the focus for the TPS implementation. TPS
would be implemented into a fifth-grade math classroom at Bedford Elementary during the 2006-2007 school year.

Throughout the summer, prior to the start of the 2006-2007 school year, Dixon worked with a group of elementary teachers to design the TPS model classroom structure and to set goals for the upcoming school year. The initial goals for the program were set at 100% of the student population meeting proficient or advanced placement on the PSSA math exam. In order to accomplish this goal, the team decided that all students would have to achieve 85% or higher on every evaluation they took. Failure to meet these goals would result in independent practice, or re-teaching the material using a different approach. The consensus of the team was to provide the appropriate differentiated instruction to enable all children to master the material. Dixon led the team to follow a process of identifying the need, specifying the work to be done, doing the work, using feedback to determine success or failure and solving problems at the root cause. The problem solving process was designed using the TPS model and followed a basic process: understand the need, define the problem, identify methods to measure success, conduct first-hand observations, conduct five why-question root cause analysis, generate ideas for countermeasures and develop the target condition, quickly implement ideas and test them, continue solving problems in real time, reflect and revise. During this summer planning, the team also designed a help line chart, problem log chart and many student-learning tools that could be readily available for student use during the 2006-2007 school year. The problem log was intended to highlight problems encountered by the classroom teacher, create a clear record of those problems, indicate the date the
problem was solved and also who solved the problem. Team members also spent large amounts of time analyzing data to learn as much as possible about each student within the target group. Clearly understanding the strengths and weaknesses of each student became the main focus for meeting student needs; as well as meeting student needs in real time.

On August 21, 2006, the core team met to finalize implementation of the Bedford Learning System project. The meeting focused on reviewing the work done to date and reviewing the goals set for the model classroom. Each member was provided a clear understanding of their role and expected time limitations to respond when help was needed. Since the project was to be funded by the Grable Foundation, time was spent reviewing the goals set forth in the grant application. Meeting dates were established for every other week throughout the implementation process and occurred on variable days during the afternoons at the Bedford Elementary School. In addition to measuring student success, the team wanted to measure changes in teacher perceptions of efficacy. To measure efficacy, the OSTES survey instrument was used prior to the start of the implementation process in August of 2006, and then again at the end of the project in June of 2007.
CHAPTER FOUR

RESULTS

This Chapter will discuss the data analysis results for three dependent variables listed in this study: 4Sight predictive benchmark assessments, PSSA scores and the OSTES. Results of the following data analysis will serve to support or reject the research hypothesis. The research hypothesis states that TPS implementation into a fifth-grade classroom will have a significant impact on student performance scores as indicated by both the PSSA and 4Sight predictive benchmark assessments. The researcher also predicted that TPS will promote a higher sense of teacher self-efficacy and that the OSTES would show statistical significance when comparing a non-TPS group to the target TPS group. Chapter Four will discuss external threats to validity, student performance data using the PSSA and then 4Sight results; teacher efficacy results will follow the student performance data. Chapters Four and Five will also include a descriptive section that will describe the implementation process and results through the experiences of the researcher, implementation teacher and the TPS consultant.

Comparison of TPS to Control Group on Demographic Variables

To limit the external threats to validity, the fifth-grade student population was studied to determine if the TPS group was similar to the control group in the areas of: gender, race, economic status, and the percentage of learning support students in each group. The TPS sample is comprised of 45 students, 23 male and 22 female students. The remainder of the fifth-grade class is comprised of 96 students, of which there are
48 males and 48 females. Gender issues threatening the validity of the results are eliminated due to the near perfect 50% make up of both the TPS group and the control group. Using the free and reduced lunch program as a guide for students who may be at an economic disadvantage, the cafeteria staff was able to break down the number of students in both the TPS group and the control group. The TPS group had 15 students receiving free or reduced lunches, while the control group had a total of 40 students in the free or reduced lunch program. Of the 45 students in the TPS group, 15 students, or 33%, were in the free or reduced lunch program. The control group was made up of 96 students, 40 students, or 41% were in the free or reduced lunch program. These data suggest that the TPS group may have had a slight advantage over the control group because a lesser percentage of the student population was enrolled in the free or reduced lunch program. Approximately three students account for the difference in economically hindered students. Special education students in the TPS group consisted of 8 students, or 18% of the population. The control group was comprised of 96 students and had a total of 15 special education students, or 16% of the sample. In this case, the TPS group may have been at a slight disadvantage because its sample group contained 2% more special education students than did the control group. As shown in these data, the TPS group and the control group were suitably balanced, with only a slight variation in economically disadvantaged students, more of whom are in the control group; however, special education students made up 2% more of the total population in the TPS group. The Bedford Area School District, for the 2006 end of the year reports, showed only 1% of its total population being of minority status. Closer examination
of the data revealed that no single category of minority status is comprised of even 1% of the total population in the district, making this statistic immeasurable and not a limitation to validity results for this study.

Teacher preparation time was another factor that may influence the validity of statistical results. Teachers in the TPS group were afforded the opportunity to work 13 additional days in the summer of 2006, which were paid days taken from the Grable grant. No other teachers in this study received pay for any work done prior to the start of the 2006-2007 school year, which could be a concluding assumption, depending on the results of the data analysis. Teachers in the control group could participate in the summer workdays if they so elected, but it would be without pay. One difference may be the process through which teachers prepare for the upcoming school year, more so than getting paid for their time.

Budget expenditures may also account for threats to the validity of results gained from this research. To limit such a threat, the budget expenditures for TPS were analyzed in detail and compared to the budget figures used by other teachers. The TPS group spent a total of $495.11 on materials that were beyond the original scope of the budget provided to other teachers. When looking at per-pupil expenditure, the TPS group spent $11.00 more on each student during the 2006-2007 school year than did the control group. Determining that budgetary limitations may influence the results of this study is very unlikely due to the limited differences in spending between the TPS group and the control group.

Though the TPS group had an established help line in place that identified each person that they could go to for help, every teacher in this study had the same
people available to them. The major difference is in the organization of the help line flow chart that TPS uses. Teachers not in the TPS study may not know exactly how to access individuals, for what reasons to ask for help, and in what order to proceed. This factor is not a limitation to validity, but part of the stimulus TPS design that should influence changes in student performance and teacher efficacy.

Teacher efficacy-results were influenced by the small sample size and may limit the researcher’s ability to make conclusive statements regarding this factor. Because of the small sample size--13 teachers--teacher efficacy results may only be assumptions that can be gleaned from the data. TPS was implemented in the fifth-grade and suggests that teacher-efficacy results fluctuated from August of 2006 until June of 2007. Adding validity to the efficacy results would be the fact that fourth-grade teachers have not been trained or exposed to TPS; therefore, the researcher suggests that the OSTES would demonstrate less gain in teacher perceptions of efficacy with fourth-grade teachers when compared to the fifth-grade teachers who have been trained in, or exposed to, TPS.

**Student performance**

The first research question to be addressed in this chapter is TPS’s relationship to student performance as indicated by PSSA student results. What is the relationship between the introduction of TPS and student math academic performance as indicated by student results from the 2006 and 2007 PSSA state exams and 4Sight predictive benchmark assessments taken during the 2006-2007 school year? The research hypothesis states that TPS will show statistical improvement in the implementation group’s math performance on PSSA test and 4Sight predictive
benchmark assessments, when compared to classrooms not implementing TPS. Of the 141 students involved in this study, 125 students took the PSSA test in both fourth and fifth-grade, while 128 students took all three 4Sight tests.

**PSSA Results**

PSSA data was collected from 125 student subjects from the original group of 141 students. The TPS group had 42 students who had taken the PSSA test in grade four during the 2005-2006 school year and grade five during the 2006-2007 school year. In the control group there were 83 students who had taken the PSSA test in grade four during the 2005-2006 school year and grade five during the 2006-2007 school year. The total number of student subjects for the PSSA section of data analysis was 125 students.

In a first analysis, the mean PSSA scores between the experimental group and the control group was compared using 2005/2006 and 2006/2007 student results. A two-way mixed analysis of variance with a between-subjects factor of group (experimental or control) and a within-subjects factor of year (2005-2006 and 2006-2007) was used to analyze the data. This analysis will allow the researcher to look specifically at differences in mean scores between the two groups and the differences within each group’s performance from 2005-2006 and 2006-2007 PSSA test results.

The first analyses compared mean PSSA scores of the experimental and control groups and for the 2005-2006 / 2006-2007 school years. A two-way mixed analysis of variance with a between-subjects factor of group (experimental or control) and a within-subjects factor of year (2005-2006 and 2006-2007) was used to analyze the data. There was not a statistically significant interaction between the school years.
and groups (2005-2006 Control $M = 1403.15$; 2006-2007 Control $M = 1420.21$; 2005-2006 Experimental $M = 1397.10$; 2006-2007 Experimental $M = 1414.14$), $F(1, 123) = 0.00, p > .999$. There was no statistically significant difference in math scores between the experimental ($M = 1405.62$) and control ($M = 1411.68$) groups, $F(1, 123) = 0.029, p = 0.865$. Finally, there was no statistically significant difference in math scores within the 2005-2006 ($M = 1400.12$) and 2006-2007 ($M = 1417.17$) school years, $F(1, 123) = 1.122, p = 0.292$. Table 1 displays the descriptive statistics for this analysis. Table 2 summarizes the results of the analysis. Figure 1 below displays the clear lack of interaction.

Table 1

**Descriptive Statistics for PSSA Analysis**

<table>
<thead>
<tr>
<th>Year</th>
<th>Group</th>
<th>$M$</th>
<th>$SD$</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005-2006</td>
<td>Control</td>
<td>1403.15</td>
<td>192.765</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>Experimental</td>
<td>1397.10</td>
<td>181.163</td>
<td>42</td>
</tr>
<tr>
<td>2006-2007</td>
<td>Control</td>
<td>1420.20</td>
<td>26.373</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>Experimental</td>
<td>1414.14</td>
<td>179.401</td>
<td>42</td>
</tr>
</tbody>
</table>

Table 2

**Analysis Results**

<table>
<thead>
<tr>
<th>Source</th>
<th>$SS$</th>
<th>$df$</th>
<th>$MS$</th>
<th>$F$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>2045.357</td>
<td>1</td>
<td>2045.357</td>
<td>0.029</td>
<td>0.865</td>
</tr>
<tr>
<td>Year</td>
<td>16221.698</td>
<td>1</td>
<td>16221.698</td>
<td>1.122</td>
<td>0.292</td>
</tr>
<tr>
<td>Interaction</td>
<td>0.002</td>
<td>1</td>
<td>0.002</td>
<td>0.001</td>
<td>0.999</td>
</tr>
<tr>
<td>Error</td>
<td>1777887.302</td>
<td>123</td>
<td>14454.368</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Of the 141 students selected for this study, 128 students participated in all three 4Sight tests and are included in this analysis. For the second analysis, there are a total of 44 students in the TPS group and 84 students in the control group. A repeated measures analysis of variance with a within-subjects factor of occasion (baseline, second quarter, and fourth quarter) and a between-subjects factor of group (experimental or control) was used to test the change in math scores. There was not a statistically significant group by occasion interaction, $F(2, 252) = 0.377, p = 0.686$, and there was not a statically significant difference between the scores of the two groups.
groups, $F(1, 126) = 0.924, p = .338$. However, there was a statistically significant change in scores between occasions, $F(2, 252) = 211.223, p < .001$.

The mean change in scores between occasions was subsequently tested using paired samples $t$-tests. Since there were three post hoc comparisons (baseline vs. second quarter, baseline vs. fourth quarter, and second quarter vs. fourth quarter), avoiding capitalization on chance findings was controlled by setting the Type I error rate at $\alpha = .05/3 = .0167$. Differences were statistically significant for all three comparisons. Specifically, 4Sight Scores increased by 5.047 from baseline to second quarter, $t(127) = 13.603, p < .001$, by 3.516 from second quarter to fourth quarter, $t(127) = 10.393, p < .001$, and by 8.563 from baseline to fourth quarter, $t(1) = 18.735, p < .001$. Table 3 displays the descriptive statistics for all of the interaction and mean effects tests. Table 4 is the ANOVA source of variance table. Table 5 displays the results from the post hoc tests. Figure 2 illustrates the mean math score change over time.

Table 3

**Descriptive Statistics**

<table>
<thead>
<tr>
<th>Group</th>
<th>Level</th>
<th>Mean</th>
<th>SE</th>
<th>95% LL</th>
<th>95% UL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>Control</td>
<td>26.155</td>
<td>0.667</td>
<td>24.834</td>
<td>27.475</td>
</tr>
<tr>
<td></td>
<td>Experimental</td>
<td>25.061</td>
<td>0.922</td>
<td>23.236</td>
<td>26.885</td>
</tr>
<tr>
<td>Occasion</td>
<td>Baseline</td>
<td>21.136</td>
<td>0.623</td>
<td>19.903</td>
<td>22.368</td>
</tr>
<tr>
<td></td>
<td>2nd Quarter</td>
<td>26.090</td>
<td>0.611</td>
<td>24.880</td>
<td>27.300</td>
</tr>
<tr>
<td></td>
<td>4th Quarter</td>
<td>29.597</td>
<td>0.617</td>
<td>28.376</td>
<td>30.818</td>
</tr>
<tr>
<td>Control Group By Occasion</td>
<td>Baseline</td>
<td>21.476</td>
<td>0.730</td>
<td>20.031</td>
<td>22.921</td>
</tr>
<tr>
<td></td>
<td>2nd Quarter</td>
<td>26.726</td>
<td>0.717</td>
<td>25.307</td>
<td>28.145</td>
</tr>
<tr>
<td></td>
<td>4th Quarter</td>
<td>30.262</td>
<td>0.723</td>
<td>28.830</td>
<td>31.394</td>
</tr>
<tr>
<td>Experimental Group By Occasion</td>
<td>Baseline</td>
<td>20.795</td>
<td>1.009</td>
<td>18.799</td>
<td>22.792</td>
</tr>
<tr>
<td></td>
<td>2nd Quarter</td>
<td>25.455</td>
<td>0.991</td>
<td>23.494</td>
<td>27.415</td>
</tr>
<tr>
<td></td>
<td>4th Quarter</td>
<td>28.932</td>
<td>1.000</td>
<td>26.954</td>
<td>30.910</td>
</tr>
</tbody>
</table>
Table 4

ANOVA Source of Variance Table

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>103.705</td>
<td>1</td>
<td>103.705</td>
<td>0.924</td>
<td>0.338</td>
</tr>
<tr>
<td>Occasion</td>
<td>4174.640</td>
<td>2</td>
<td>2087.320</td>
<td>211.223</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Interaction</td>
<td>7.452</td>
<td>2</td>
<td>3.726</td>
<td>0.377</td>
<td>0.686</td>
</tr>
<tr>
<td>Error</td>
<td>2490.277</td>
<td>252</td>
<td>9.882</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5

Post Hoc Test Results

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Mean 1</th>
<th>Mean 2</th>
<th>Difference</th>
<th>SE</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline – 2nd Qtr</td>
<td>21.24</td>
<td>26.29</td>
<td>5.05</td>
<td>0.371</td>
<td>13.603</td>
<td>127 &lt;.001</td>
<td></td>
</tr>
<tr>
<td>2nd Qtr – 4th Qtr</td>
<td>26.29</td>
<td>29.80</td>
<td>3.51</td>
<td>0.338</td>
<td>4.185</td>
<td>127 &lt;.001</td>
<td></td>
</tr>
<tr>
<td>Baseline – 4th Qtr</td>
<td>21.24</td>
<td>29.80</td>
<td>8.56</td>
<td>0.457</td>
<td>7.658</td>
<td>127 &lt;.001</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2

Group

---

Control

---

Experimental
Teacher Efficacy

The final analysis for this study compares two groups of teachers (fourth-and fifth-grade) using the OSTES to relate changes in teacher’s perceptions of efficacy. There are a total of thirteen teachers that took the pre-and post-surveys; of those, six teachers are from grade four and seven are from grade five. This analysis addresses the following research question: What is the relationship between the introduction of TPS and teacher efficacy of fourth-and fifth-grade teachers as indicated by the Ohio State Teacher Efficacy Scale during the 2006-2007 school year?

The next set of analyses pertains to the efficacy data. Reliability, based on Cronbach’s alpha, was deemed to be more than at 0.895 for the 24 question OSTES survey. There was not a statistically significant mean score difference between the groups. The lack of significance between group mean scores is based only on descriptive data. Inferential statistics were not used because the sample size was too small to make conclusions about the data. Table 6 displays the descriptive statistics for this analysis. Table 6 also suggests a marked interaction where fifth-grade efficacy scores are higher than fourth-grade efficacy scores at pretest. Then, at posttest, the fifth-grade efficacy scores increase while the fourth-grade efficacy scores appear to actually decrease.
Table 6

**Descriptive Statistics**

<table>
<thead>
<tr>
<th>Group</th>
<th>Occasion</th>
<th>M</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td>4th Grade</td>
<td>169.50</td>
<td>12.97</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>5th Grade</td>
<td>173.14</td>
<td>10.88</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>171.46</td>
<td>11.53</td>
<td>13</td>
</tr>
<tr>
<td>Posttest</td>
<td>4th Grade</td>
<td>167.83</td>
<td>19.62</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>5th Grade</td>
<td>177.86</td>
<td>15.93</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>173.23</td>
<td>17.73</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 6 above shows a convergent trend that demonstrates an increase in fifth-grade teacher efficacy scores at posttest; while at the same time a decrease in fourth-grade teacher efficacy scores at posttest. Despite the trend towards greater levels of teacher efficacy for fifth-grade teachers, there was no significance found in the statistics, possibly due to the small n. The following item analysis shows the actual groups mean scores, from both the fourth and fifth-grade teachers, for each question on the survey, Appendix G. A closer look at the question analysis below indicates five questions on the OSTES survey where fifth-grade teachers had a lower group mean score at posttest than did the fourth-grade teacher group. Those five items are marked in bold and are the following questions: question 3, question 6, question 7, question 9 and question 13.

Table 7

**OSTES Group Mean Scores**

<table>
<thead>
<tr>
<th>Question</th>
<th>4th Pretest</th>
<th>4th Posttest</th>
<th>5th Pretest</th>
<th>5th Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 1</td>
<td>6.33</td>
<td>6.50</td>
<td>6.86</td>
<td>7.43</td>
</tr>
<tr>
<td>Question 2</td>
<td>6.33</td>
<td>7.16</td>
<td>7.14</td>
<td>7.71</td>
</tr>
<tr>
<td>Question 3</td>
<td>7.83</td>
<td>7.50</td>
<td>7.43</td>
<td>7.43</td>
</tr>
<tr>
<td>Question 4</td>
<td>7.00</td>
<td>6.66</td>
<td>6.57</td>
<td>6.86</td>
</tr>
<tr>
<td>Question</td>
<td>Fifth-Grade</td>
<td>Fourth-Grade</td>
<td>Fifth-Grade</td>
<td>Fourth-Grade</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>--------------</td>
<td>-------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Question 5</td>
<td>8.66</td>
<td>8.50</td>
<td>8.57</td>
<td>8.57</td>
</tr>
<tr>
<td>Question 6</td>
<td>8.50</td>
<td>8.16</td>
<td>7.86</td>
<td><strong>7.86</strong></td>
</tr>
<tr>
<td>Question 7</td>
<td>7.66</td>
<td>7.83</td>
<td>7.29</td>
<td><strong>7.29</strong></td>
</tr>
<tr>
<td>Question 8</td>
<td>8.50</td>
<td>8.16</td>
<td>8.71</td>
<td>8.57</td>
</tr>
<tr>
<td>Question 9</td>
<td>7.83</td>
<td>8.16</td>
<td>7.28</td>
<td><strong>7.57</strong></td>
</tr>
<tr>
<td>Question 10</td>
<td>7.00</td>
<td>7.66</td>
<td>7.86</td>
<td>8.14</td>
</tr>
<tr>
<td>Question 11</td>
<td>6.83</td>
<td>6.83</td>
<td>6.86</td>
<td>7.43</td>
</tr>
<tr>
<td>Question 12</td>
<td>6.66</td>
<td>7.50</td>
<td>7.14</td>
<td>7.71</td>
</tr>
<tr>
<td>Question 13</td>
<td>8.00</td>
<td>8.33</td>
<td>8.29</td>
<td><strong>8.29</strong></td>
</tr>
<tr>
<td>Question 14</td>
<td>6.83</td>
<td>6.16</td>
<td>6.71</td>
<td>7.43</td>
</tr>
<tr>
<td>Question 15</td>
<td>7.83</td>
<td>7.66</td>
<td>7.00</td>
<td>7.71</td>
</tr>
<tr>
<td>Question 16</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Question 17</td>
<td>6.50</td>
<td>6.50</td>
<td>7.29</td>
<td>7.43</td>
</tr>
<tr>
<td>Question 18</td>
<td>7.16</td>
<td>7.16</td>
<td>8.00</td>
<td>7.71</td>
</tr>
<tr>
<td>Question 19</td>
<td>7.33</td>
<td>7.33</td>
<td>7.00</td>
<td>7.71</td>
</tr>
<tr>
<td>Question 20</td>
<td>7.66</td>
<td>7.33</td>
<td>6.86</td>
<td>8.00</td>
</tr>
<tr>
<td>Question 21</td>
<td>7.66</td>
<td>6.83</td>
<td>7.00</td>
<td>7.43</td>
</tr>
<tr>
<td>Question 22</td>
<td>7.00</td>
<td>6.66</td>
<td>6.86</td>
<td>7.43</td>
</tr>
<tr>
<td>Question 23</td>
<td>7.16</td>
<td>7.16</td>
<td>7.86</td>
<td>8.00</td>
</tr>
<tr>
<td>Question 24</td>
<td>7.00</td>
<td>7.66</td>
<td>8.43</td>
<td>7.86</td>
</tr>
</tbody>
</table>

The OSTES survey instrument is designed around three dimensions of efficacy: instructional strategies, student engagement and classroom management. The survey also includes questions that draw upon the following teacher efficacy dimensional areas: modeling, mastery experiences, persuasion and physiological arousal. When we take a closer look at each question on the survey that the fifth-grade teachers scored lower on than did the fourth-grade teachers, we find that those questions do not deal specifically with instructional strategies or student engagement.
The focal point of the TPS classroom model design was based on instructional techniques and process flow to improve student learning; which would include teacher efficacy categories of instructional strategies, student engagement and possibly mastery experiences. The following questions yielded a lower group mean score for fifth-grade teachers.

3. How much can you do to control disruptive behavior in the classroom?
6. How much can you do to get students to believe they can do well in schoolwork?
7. How well can you respond to difficult questions from your students?
9. How much can you do to help your student’s value learning?
13. How much can you do to get children to follow classroom rules?

Question number 3 focuses on classroom management procedures that would not be enhanced by the TPS model classroom design. Classroom management may improve as a result of better instructional processes, but most likely would not be any more than teachers that were not exposed to the TPS process. Question number 6 focuses on persuasion and physiological arousal, which may have been enhanced by TPS, but was not a targeted outcome. Question number 7 focuses on classroom management and mastery experiences; teachers in the TPS model classroom would not be expected to score higher in this area. Question 9 focuses on physiological arousal and would not be a targeted outcome for the TPS implementation. Finally, question number 13 focuses on classroom management, which would not be enhanced by the TPS implementation. After taking a closer look at the five questions where the fourth-grade teachers scored higher, it is clear that those questions did not
deal specifically with instructional strategies or student engagement and therefore may not be enhanced by TPS.

The teachers in this survey did not answer question number 16 because of a survey error on the pre-survey instrument, leaving a total of 23 questions to be answered. Of the 23 questions answered, fifth-graded teachers scored higher than fourth-grade teachers on 18 questions. Most of the 18 questions that the fifth-grade teachers scored higher on included efficacy dimensions of instructional strategies and student engagement. There were eleven questions that focused primarily on instruction and student engagement. Fifth-grade teachers scored higher than fourth-grade teachers on each question that focused primarily on instructional practices and student engagement. Those questions are listed below.

1. How much can you do to get through to the most difficult students?
2. How much can you do to help your students think critically?
10. How much can you gauge student comprehension of what you have taught?
11. To what extent can you craft good questions for your students?
14. How much can you do to improve the understanding of a student who is failing?
17. How much can you do to adjust your lessons to the proper level for individual students?
18. How much can you use a variety of assessment strategies?
19. How well can you keep a few problem students from ruining your entire lesson?
22. How much can you assist families in helping their children do well in school?
23. How well can you implement alternative strategies in your classroom?
24. How well can you provide appropriate challenges for very capable students?
The questions listed above clearly focus on dimensions of teacher efficacy that would be directly impacted by the TPS implementation project. On each of these questions the fifth-grade teachers scored higher than their peers at the fourth-grade level. This would be an expected outcome from the TPS implementation. There are seven questions that did not deal specifically with instructional strategies or student engagement that the fifth grade teacher’s answers resulted in a higher mean score than the fourth-grade teachers. Those seven questions are listed below.

4. How much can you do to motivate students who show low interest in schoolwork?

5. To what extent can you make your expectations clear about student behavior?

8. How well can you establish routines to keep activities running smoothly?

12. How much can you do to foster student creativity?

15. How much can you do to calm a student who is disruptive or noisy?

20. To what extent can you provide an alternative explanation or example when students are confused?

21. How well can you respond to defiant students?

After reviewing the questions on this survey, it is clear that the fifth-grade teachers did show higher group mean scores on questions that focused primarily on instructional strategies and student engagement. The seven questions listed above indicate that the fifth-grade staff also showed higher mean scores for efficacy dimensions of persuasion, modeling, physiological arousal, and mastery experiences; not in all cases though. There were six questions on this survey that focused specifically on classroom management skills; of those six questions the fifth-grade
teachers scored higher than the fourth-grade teachers on three. When looking at only those questions dealing with instructional strategies and student engagement, the fifth-grade teachers had higher group mean scores for 100% of those questions. The TPS model classroom had an impact on those questions that specifically addressed instruction and student engagement.

Classroom Teacher Interview

The descriptive part of this study involved interviewing the TPS implementation teacher using a five-question document, Appendix I, to prompt responses. Interview questions followed the common themes represented in the research questions for this study. It was the intention of the researcher to include an interview with the classroom teacher that would allow readers to gain a perspective from the implementation level. Questions were designed around the basic concepts of student achievement, teacher efficacy and the actual instructional differences that the classroom instructor noticed while implementing TPS. The first question focused on any differences that the instructor noticed between TPS and her previous traditional classroom practices. The second question focused on the classroom teacher’s perceptions about her own professional growth throughout the implementation process. This question was designed to gain further insight into the results of the teacher efficacy study. Because there was only one teacher trained in TPS, it is essential to hear the teacher’s perception of her own professional growth. The third question was designed to look deeper into student performance by analyzing the classroom teacher’s perception of student academic growth, student attitudes towards learning, and parental feedback that was shared during the course of the
implementation project. Question number four was open ended to allow the implementation teacher an opportunity to communicate any major eye opening experiences that will permanently change her future classroom teaching practices. The fifth and final question was left completely open ended and focused on any findings that may not have been discussed in the previous interview questions. It was the intention of the researcher to limit the interview time to approximately 45 minutes. Questions asked of the TPS teacher were designed to gain further insight into the actual impact that TPS had on classroom instruction from the teacher’s perspective.

Prior to the interview, the classroom teacher and Dixon both had an opportunity to review the questions and discuss findings. In June of 2007, the researcher met with the classroom teacher at Bedford High School to complete the interview. Copious notes were taken while the interview was being recorded to ensure that the classroom teacher’s actual insights could be accurately described. Following the interview, the researcher submitted a copy of the questions to the classroom teacher in electronic form. The purpose of the electronic form was to allow the TPS instructor an opportunity to add anything she may have missed in the interview. Immediately after the interview process was complete, the researcher carefully wrote up the responses to each of the five questions and resubmitted them to the TPS classroom teacher to insure accuracy. The following questions are the actual interview questions and the responses are an accurate summary of the TPS teacher responses.
**Question One:** What are the major differences between the TPS model and a traditional classroom model from the teacher’s perspective?

The main difference is that there is a goal of 85% achievement for all students; if that goal is not met, re-teaching or remediation occurs. What gets taught is based on student needs, and the student needs are determined by any of the following: classroom assignments, PSSA scores, or 4Sight scores. The classroom is much more data-driven than before, but for this to work effectively, the data must be readily available for the teacher. Time does not exist for the teacher to search out information.

There is a constant expectation of 85% and if students are not achieving at that level, the teacher must find another way to teach the material. The 85% may be a bit too high for all students to reach, but there were only four students that did not make the 85% and only one that did not reach the 80% mark. In the past, with child study, teachers would look for ways to get the students to pass at the 65% level. That did not work because students were moving to the next grade without the necessary skills. Students were passing, but they did not have the background knowledge they needed to be successful in the next grade level. The high achieving students are still performing well; however, in the TPS model students do not receive instruction on information they have already mastered.

**Question Two:** Looking back over your experience implementing TPS, how would you describe the impact it had on your professional growth as a teacher?

TPS has created a necessity for the teacher to focus on individual students based on their needs within the curriculum. Material is no longer covered just because
it is in the curriculum. Without an 85% mastery level from all students, the material is re-taught in a different manner until 85% is achieved. TPS encourages the instructor to keep asking why, until the root cause of the problem can surface. Many times the students are able to answer their own questions and if they cannot, asking why long enough tends to surface the problem. Before TPS, some students would not achieve at the 85% level, but would not ever ask why. The problem solving aspect of TPS has brought about a completely different view of the teaching/learning process for this particular instructor. Major growth has occurred in the area of student questioning. No longer are answers given to students. The process of leading students to the correct answers through questioning tends to create a more thorough learning process.

Perhaps the most important change indicated by the TPS instructor is in the self-efficacy portion, believing that if she can match her methods of inquiry to the students needs, they will understand the material. No longer is it acceptable to pass a student along without their mastering the material. The earlier this process starts the better chance schools have of meeting the needs of all learners. It is also equally important that the parents get involved, especially with students who are struggling.

*Question Three: How would you describe the impact that TPS had on student performance, student perceptions, and parent perceptions and on other fifth-grade staff members?*

Students are math thinkers, not just doers of the algorithm. They have greater concept knowledge. The shift to inquiry method meant not just telling students the way to do it or the answer. This method seemed to increase their confidence with math. Students like math because we are using a variety of methods to gain
understanding. They know the goal is 85% and will receive the necessary help to achieve at the 85% level. Very few parents contacted the instructor throughout the 2006-2007 school year regarding math questions. Students were willing to work on areas of weakness within the curriculum, as long as the help was available. Throughout this implementation, one thing stood out as a message from students; they perceived that the instructor cared about their learning.

In general, the instructor perceived the parents to be happy with the math instruction and with their child’s progress. The instructor thinks that parents will most often refrain from contact when they are satisfied with the instruction their child is receiving. There were no parent complaints voiced to the teacher about math instruction. Staff members did seem to be skeptical of the 85% achievement requirement, but most were very willing to share ideas to help acquire the goal. Many of the staff members were concerned about the additional workload that TPS may put on them and were reluctant to volunteer for the implementation. Teachers in the fifth-grade were supportive of the TPS effort and shared ideas, but most teachers took the role of the observer.

Question Four: Are there any major eye-opening experiences that may impact the methods of instruction you choose to practice as a result of this exposure to TPS?

According to the TPS implementation teacher, the value of peer interaction is tremendous. Meeting with the TPS consultant and other teachers on a bi-weekly basis, proved that solving problems was much more efficient and effective with a group of concerned people. The helpline proved to be invaluable; being able to call someone for help, as soon as it is needed, was a new practice for this teacher. Asking
the question of “why” the problem is occurring is critical, instead of just making assumptions for the problem existing. This new method of problem solving will be an embedded philosophical shift in practice, according to the TPS instructor, all students can learn and achieve if the teaching matches their needs. In the past, when a student missed four questions on a quiz, the teacher would move on to the next topic; under the TPS umbrella, the four questions are examined to see if the mistakes are consistent to any lack of instruction.

*Question Five: Do you have any other significant findings, as a result of your experience with TPS, which may be of interest to this study or the educational research community?*

If educators would start the TPS process at the kindergarten level and use the 85% goal and the problem solving approach, they should produce more proficient learners. The biggest frustration noted by the TPS teacher was the four students who consistently did not reach the 85% goal due to their learning gaps from previous grades. She noted that if students missed learning when the goal was 65%, and over the years it compounded, by the time they were in the fifth-grade, it was almost impossible to catch them up.

*Research Findings*

The first analysis focused on student PSSA scores and was designed to test the hypothesis that the TPS group’s fifth-grade results would increase over their fourth-grade results, when compared to the control group. A two-way mixed analysis of variance with a between-subjects factor of group and a within-subjects factor of year was used to analyze the data. There was clearly no significance in the results that
could allow one to conclude that the TPS group’s mean performance score was greater than the mean score for the control group.

A second analysis was designed to test the hypothesis that the TPS group’s mean scores on 4Sight Predictive Benchmark Assessments would be higher than the control group’s results. To test this hypothesis, the researcher used a repeated measures analysis of variance with a within-subjects factor of occasion and a between-subjects factor of group to test the changes in 4Sight math scores. There was not a statistically significant difference between the groups by occasion interaction, nor a difference between the TPS group and the control group. The mean change in scores between occasions was tested using a paired samples t-test. Results showed that there was a statistically significant difference in all three post hoc comparisons: baseline vs. second quarter, baseline vs. fourth quarter, and second quarter vs. fourth quarter. The data does not suggest that the TPS group performed better than the control group, but it does indicate that both groups showed considerable improvement throughout the year.

In a third and final analysis, the researcher tested the hypothesis that teacher efficacy would improve, as indicated by the OSTES; within the fifth-grade staff when compared to the survey results from the fourth-grade teaching staff. There was not a significant mean difference between group mean scores from pre to post surveys. Though there was not a significant statistical difference indicated by this experiment, the results suggested that a larger sample size was needed to draw conclusive results. For example, the fifth-grade teaching staff recorded efficacy scores in the initial survey that were much higher than that of fourth-grade teachers. When the post
survey results were compiled, the fifth-grade teacher scores increase, while the fourth-grade teacher scores actually decrease. This could be a result of the small sample size and will be discussed in Chapter Five.

Careful analysis of the OSTES survey instrument indicates that the fifth-grade teachers did show higher group mean scores for all questions that related directly to instruction and student engagement. The only questions that the fourth-grade teachers showed higher mean scores for were those related to classroom management and mastery experiences. As noted earlier in this chapter, the TPS implementation focused primarily on improving student performance through an innovative classroom design; therefore, it would not be expected that the fifth-grade teachers would score higher or lower than their peers at the fourth-grade level on questions regarding classroom management. That was exactly the case as indicated by the descriptive analysis of the OSTES survey.

In a final attempt to capture any major findings that the TPS implementation project had on teacher efficacy or student learning, the researcher used a five-question document to interview the classroom teacher. The results of this interview clearly indicate that the teacher perceived noteworthy growth in her own professional practice, in her student’s understanding of math concepts, and in her belief that all students can achieve at the 85% level if the instruction is matched to the student’s needs. Chapter Five will discuss in-depth the implications of the descriptive results for this study.
CHAPTER FIVE
SUMMARY OF MAJOR FINDINGS

Chapter Five will discuss the statistical research findings and any major learning that were gained as a result of the TPS implementation project at Bedford Elementary School. This case study was designed to test the effectiveness of the TPS management system to improve student math test scores and to increase teacher efficacy.

The contents of this chapter will include a brief summary of the research findings, related learning from the classroom teacher interview, implications and recommendations for further studies. In addition to the overview, the author will discuss interpretations and related educational findings that may influence future educational practices and studies.

Statement of the Problem

It is no secret to anyone associated with public education that accountability for student learning is here to stay. NCLB was passed in an effort to ensure that all students achieve at a minimum standard level or above. With the growing pressures placed upon public education to meet the needs of all learners, it has become increasingly apparent that public schools will have to change with the demands of its customers. Schools are now held accountable for student learning and with each passing year, the acceptable levels of achievement are being raised. Today’s society has become accustomed to testing in most avenues of life; it is no surprise that education will continue to be a focal point for high-stakes testing and standards.
With the growing demands placed upon public schools and the continued pressures of legislation efforts to improve student learning, schools will be forced to move out of the norm and search for better ways to conduct the business of education. Public school management and operations are becoming a hot topic for educational research. Some believe learning can be done more efficiently and effectively using models outside the educational world. Businesses that have shown momentous gains during their years of operation and have survived economic disasters are focal points for future educational research. The time is now for educational reform. The future success of public education will be determined by a leader’s ability to effectively meet the increasing demands of an ever-changing society. This case study represents an initial attempt to implement a future focused industrial model into a fifth-grade classroom at Bedford Elementary School.

When summarizing an implementation project of this magnitude, the total scope of learning cannot be measured entirely through statistical analysis; much of the valuable learning that has been gained from such a future-focused project will come from discussions with those closest to the process. In order to capture any relevant information gleaned from this case study, the researcher chose to include a descriptive component to describe the process from the teacher’s point of view. The following discussion will be based entirely on the statistical data and results that can be concluded from that data. Other important information gained from the TPS implementation will also be discussed in further detail.
Discussion

In this case study, there were 3 dependent variables used to measure changes in the target and control group’s performance: student PSSA scores, student 4Sight Benchmark Assessment scores and the OSTES survey to measure teacher efficacy. The data results for student performance consisted of 2 dependent variables being used to measure mean changes in performance between the TPS group and the control group. PSSA results indicated that there was no statistically significant interaction between school years and groups. There was also no conclusive data that would support the hypothesis that the TPS group would score higher than the control group on the PSSA exam.

Data results for the 4Sight Benchmark Assessments indicated statistically significant changes in student performance for both the TPS group and the control group when comparing mean changes between occasions. This data would indicate that both the experimental group and control group showed statistically significant gains in student performance, but neither group scored higher than the other. There was not a statistically significant group by occasion interaction, nor a significant difference between scores for the 2 groups.

Teacher efficacy data revealed that there were no statistically significant changes that could be documented between the 2 groups using mathematical comparisons. There was not a statistically significant mean difference between the group’s mean scores. Inferential statistics could not be used due to the small sample size and limitations placed on conclusive data. Efficacy results were based only on descriptive statistics.
The classroom teacher interview proved to be a valuable tool for gaining further insights into the TPS implementation process and to the conclusions for this research project. As discussed in Chapter Four, the classroom teacher indicated that the TPS project allowed for the use of data to effectively modify instructional practices and to gain consistency in her teaching approach. It was also clear that the classroom teacher saw her own professional growth to be impacted by the TPS process and that her own perceptions of self-efficacy were increased. According to the classroom teacher, parents and students supported the TPS process and students actually placed a higher value on learning because they saw that the teacher would not be satisfied with less than 85%.

**TPS Implementation Discussion**

The first barrier encountered by the team was the fact that PSSA results were not available until two weeks before the start of school. To gain an understanding of student needs, the team decided to use a pre-test before each chapter and rely on 4Sight test results to help direct instruction. August 29, 2006 marked the first day of school for the 2006-2007 school year and the start of the TPS project in two selected fifth-grade classrooms. Mastering Math Facts pre-tests were used during the first week of school to determine student needs and to allow for selective placement into the individualized math facts program. In addition to the Mastering Math Facts pre-tests, the instructor utilized Fast Math and Compass Learning programs with students who were struggling with math facts fluency. The TPS team saw it as essential that students have a firm grasp of math facts.
During the first two months of school, the team followed a plan to have students take a pre-test prior to the start of each unit. The purpose of the pre-test was to determine if students had already mastered the skills necessary for 85% mastery, or if additional instruction was necessary. In November of 2006, the instructor and the TPS team decided that the pre-tests were consuming too much instruction time and the students were performing poorly on the tests with new material. To maximize instruction time, the instructor decided to abandon the pre-test process and focus on the group with initial instruction. It was now a problem of how to bring instruction to the individual level, using data to guide the process.

The TPS team decided in late fall to make another attempt to understand the ‘current condition’ for each child. It was problematic for the instructor and the team, as the discussion led them once more toward using a pre-test design. Knowing that their first attempt with this process was unsuccessful, even almost wasteful, the team knew that their approach had to be different from what was previously attempted. The instructor decided to begin using the data to understand each student’s strengths and weaknesses according to specific skills and not focus on aggregate data. Students who performed poorly on a measurement unit only needed instruction with the metric portion of the test; therefore, the instruction for those students should focus primarily on the use of the metric system of measurement. In another case, students may be struggling with only the volume portion of a measurement test and their instruction could be narrowed down to focus only on volume. It was at this point that the TPS team began to see the advantage of data and how it could be applied through the TPS model to improve student learning. By the start of the second semester, the TPS
classroom was once again using pre-test activities to start each chapter, but now, the data was being analyzed for each student’s specific learning needs.

Analyzing data at the individual level created some new problems, one of the most obvious being a general lack of time. The instructor found it essential to have a data analysis specialist readily available to sort the data and provide constant feedback to guide instruction. Having an additional person to sort the data is something new for most teachers. In the TPS model classroom, it became clear to the instructor that the only way to maintain the original process was to have the additional data support person available for reassessments.

As the spring semester evolved, the TPS team grew stronger in their understanding of the 4Sight Benchmark Assessment results and how they could be used to guide instruction at the student level. In the day-to-day grind of teaching, instructors do not have the time available to analyze various forms of assessment and modify instruction to meet those needs. The TPS instructor saw that instruction could be narrowed down to the individual level if the data was readily available, in “real time”. The 4Sight data was used to assign students to individual learning activities that would focus on their areas of weakness. Group instruction continued to be a method of delivery, especially when a new concept was being introduced, or if the majority of the students tested poorly on a particular skill.

The TPS instructor found 4Sight to be helpful, but due to the small number of questions, it was difficult to focus on any one skill. 4Sight was designed to be a predictive test, to help teachers identify students who may perform poorly on the PSSA. Using measurement as an example, a student may score poorly on that
particular section of the test, but it was unclear what part of measurement they had mastered and which skills needed further instruction. For this reason, the instructor found that local classroom assessments might be the best source of data, taking into account that the assessments are aligned to the anchors. The team came to the conclusion that a student would have to take the 4Sight test three times to determine if a student had the skills necessary to master each anchor at the 85% level. If a student answered a particular question correctly only one time out of three, the team determined that additional instruction was necessary.

The pilot year for TPS was a learning experience for all those involved and according to the classroom teacher, it changed her teaching methods and her philosophy of educating students. If the teaching strategy focused on the learning needs for each student, then each student was capable of mastering the material. Another strategy that grew out of the TPS pilot was the use of questioning to assess knowledge of math concepts. The TPS instructor began using the training she received by a math consultant to help students gain an even deeper understanding of math concepts. Dr. Dan Miller provided the training and the focus was on students engaging in “accountable talk”. Early into the process, the TPS teacher noticed that students could assess their own learning needs using a variety of peer interactions and discussions.

Nearing the end of the implementation year, the team decided that they understood how to use the data much better and that the data could be narrowed down to the individual student skills. Additional support systems were put in place to help alleviate the need for instruction in many different areas. After-school tutoring was
The Toyota Way

guided by data to meet student needs, along with the addition of a computerized
version of Compass Learning for individual practice in specific skills areas. The
demands of meeting every student’s needs on an individual level became quite a task,
according to the classroom teacher, but she would never revert to her previous
practice.

Conclusions

The statistical results gained from this study indicate that the TPS model
classroom design performed well in the model year implementation and the students
made considerable gains throughout the year in their understanding of math concepts.
Due to the small sample size and the grade level selected for this study, it is difficult
to conclude that the TPS group performed better than the control group. Actually,
students within the control group performed slightly higher on 4Sight than students in
the TPS group. Statistically, there was no evidence to support that either group
performed better on the PSSA test or on the 4Sight Benchmark Assessments.

Results for the teacher efficacy component of the study yielded similar data;
there were no significant statistical changes in teacher’s perceptions of efficacy
between the fourth and the fifth-grade teachers. This portion of the study was limited
due to the small sample size of teachers. There were only thirteen teachers who
participated in the teacher efficacy study; therefore, inferential statistics were not
used. Of those thirteen teachers, only one fifth-grade teacher actually received
training in TPS. Through interactions at team meetings, faculty meetings and other
grade level events, the researcher predicted that the fifth-grade staff would show an
increase in efficacy scores. There was a notable trend showing improvement for the fifth-grade teacher efficacy scores; however, the trend was not statistically significant.

Several factors in the teacher efficacy data could lead to speculation that a larger sample of trained teachers may yield a different result. Specifically, the fifth-grade teachers in this study scored higher than the fourth-grade teachers in their initial pre-test survey. Post-survey results indicate that the fifth-grade teacher’s perceptions of efficacy increased, while the fourth-grade teacher’s perceptions of efficacy actually decreased. This is a notable change in the teacher’s perceptions of efficacy that could not be substantiated in this study, possibly due to the small sample size and limited number of trained teachers.

**Implications**

If this study would have yielded a p-value of .03 with a random group sample, we could conclude that our results would be similar 97 times out of 100; alternatively, we would make the wrong decisions about our results 3 times out of 100. There are many ways to conceptualize this and the validity of the explanation. The bottom line is that the non-significant results say absolutely nothing about “significance.” The results speak to the certainty given that there is only one sample; the p-value does not tell us anything about big or small differences. A three dollar-per-year increase in salary means nothing to most people, while a three thousand dollar-per-year increase is a different story. If we just toss out the number three, without the details, it is hard to decide if three is big or small. The p-value does not provide that information; it just helps us decide if we can trust the number three. The moral of the story is that “non-significant” results do not necessarily equal “not important” results. This can be
viewed as statistical versus practical significance. It is also noted that the number of participants in the sample group limited this study; a larger sample may have concluded a different p-value. The small sample size was listed as a limitation in this study, and it very well may be a factor in the end result of a non-significant difference in student performance.

Student performance results for this study did not produce significance for the PSSA tests or 4Sight Benchmark Assessments. What the data did show us is that both the control group and the experimental group showed statistically significant gains in student performance between each 4Sight test given. We could conclude that the TPS classroom model performed just as well as traditional classroom techniques, possibly with more predictability, due to the data-informed decision-making process used in TPS. TPS offered this classroom instructor a scientific approach to solving problems and for structuring student lessons. By utilizing data results to guide instruction, the teacher has created a working system that will support continuous improvement of instructional practice.

Another factor that has affected student performance on the PSSA and 4Sight tests is the fact that all of the students in this study had five years of previous knowledge that was not based on TPS strategies. It is entirely possible, given the small sample size, that the students in the TPS group did not have the same base knowledge level to begin the year. This was actually noted in the interview by the TPS implementation teacher. Though no factual information can be presented to argue this position, TPS implementation at lower grade levels would help to eliminate
the possibility of students starting the year with such a wide range of differing abilities.

Teacher efficacy was noted earlier in the literature review, as one of the critical components in determining if a new technique was successful or not. When we look only at the statistical data provided by the numbers, one can only conclude that the results were not significant. If we look closer at the data, we find that the fifth-grade teachers scored higher on the pre-test and the fourth-grade teachers. This could mean that before this study began, the fifth-grade teaching staff had a higher sense of self-efficacy than their peers at the fourth-grade level. To complicate the results further, the data indicated that the fifth-grade teaching staff showed an increase in teacher efficacy following the TPS implementation, while the fourth-grade teachers actually showed a decrease in their perceptions of efficacy. Given a larger sample size of teachers, the results may be much different. It is also noted that the statistical analysis was used to compare group mean scores and only one fifth-grade teacher was actually trained in TPS strategies. It was the researcher’s position that the fifth-grade group would gain valuable insights from the TPS teacher through team meetings, staff meetings and casual interactions since they work together all year. The interview with the TPS teacher would allow one to speculate that her perceptions of self-efficacy increased; but casual interactions may not have affected the remainder of the team as thought by the researcher. Teacher efficacy would be better tested with a larger number of trained teachers and a larger sample.

In addition to the above results for teacher efficacy, the survey analysis indicated important findings that support the idea that TPS will improve teacher
efficacy. Fifth-grade teachers did show higher group mean scores for each question that related directly to instruction, engagement and even the use of assessments to guide instruction. This is important because the TPS model classroom was designed to improve instruction through a systematic process that utilized assessment data to guide instruction. The TPS model classroom did not focus at all on classroom management or persuasion. It is the opinion of this researcher, according to the data, teacher interview and the question analysis, that TPS did in fact increase teacher efficacy among the fifth-grade teachers in the targeted areas.

One of the major differences between the TPS model classroom and a traditional classroom, according to the TPS teacher, was the expectation that every student should achieve at the 85% level, or higher, on every assessment. In the interview, the teacher stated that she would have targeted 65% for the lowest performing students in the past. In essence, students would be passed on to the next grade-level without the necessary skills needed to succeed. It was also the opinion of the classroom teacher that her instructional practices were being guided by student assessments.

It can be concluded from this study that students do respond to high expectations. It can be assumed from the teacher interview results that if instruction is tailored to meet the specific needs of each student, they are all capable of learning. Equally important is the increased confidence that the TPS teacher has in her ability to influence student learning.

The TPS model classroom focused on the use of the scientific method to surface and solves problems at their root cause. By using a standard method to solve
problems and analyze student performance, the TPS classroom should be capable of producing more consistent results for each student. If the teacher bases the instructional needs of each student on relevant and accurate data, then the decision-making process should be more consistent with providing appropriate instruction. During this first implementation of TPS, the data results did not show that the TPS group performed higher than the control group; but the data did indicate that both groups increased their scores significantly each time students were tested. Using a standardized system where teachers can learn from previous work should allow for an environment of continuous improvement to take place. Employees within the TPS design are expected to work smarter as a result of their experiences. In education we often change the work design so rapidly that teachers cannot use previous experiences to improve instructional design. TPS does offer education a new approach that may provide a process for which teachers can become more effective through their experiences.

The TPS model encourages teachers to solve their own problems, and to follow those problems to the root cause. This is a new concept for education; often teachers work in isolation from each other and really do not understand the root causes of many situations. TPS offers education a different look at the problem-solving process, but it does not come without a commitment. Teachers in the TPS model must be dedicated to continuous improvement and believe that all students can learn. It was the opinion of the TPS instructor that by matching the instruction to the specific needs, students will learn. Often the specific needs of each learner cannot be discovered without following the problem-solving process to the root cause. TPS
The Toyota Way encourages the “five why” process, which basically means to ask why at least five times until one understands exactly what the problem is. If a student is performing below the expected 85% level, the teacher must ask why. It may be possible that the student is not completing their homework assignments; again we must ask why. Once it is discovered that the student is missing too much school, the teacher must ask why yet again. Though this process takes more time than just lowering the standard to 65%, in the end, the teacher will understand the needs of the student much better than they would have if they did not ask why.

The help line was established to gain immediate attention to problems that could not be solved by the classroom teacher. This was a completely new concept to the classroom teacher, as her previous experiences taught her to only ask for help in extreme situations. In the TPS model, the teacher should ask for help when students are not meeting their target goals and when she cannot change the student’s level of performance on her own. What a remarkable concept. If educating students is the business that education does, why not make each student’s education the priority?

Education is the product of schooling; it is the responsibility of the system to provide the best opportunities possible to produce a quality product every time. This is a major philosophical change for education, and in the opinion of this researcher, it may be the single most important learning experience gained from this study. Educators must make the priority of student learning, their priority.

Recommendations

Due to the credible feedback gleaned from the interview with the TPS implementation teacher and the fact that the TPS students gained valuable learning,
according to the data, at least equal to their peers; it would be advised to pilot TPS again at lower grade levels. TPS would be most appropriate to pilot again at grades such as kindergarten, grade one and two. The rationale for recommending a change in grade level selection for a follow-up study is to eliminate the issue of previous knowledge. PSSA results and 4Sight Benchmark Assessments should provide an accurate measure of student performance using mean score differences, as long as the sample of students is large enough to make a statistical determination. Teacher efficacy results can be accurately measured using the OSTES, but a larger sample of instructors is necessary, and more teachers should be trained in TPS prior to the implementation. Future attempts to measure teacher efficacy would be better suited if the OSTES instrument was modified, or if another instrument was available that focused on those specific targeted areas that TPS is designed to improve. Selecting an instrument that focused on instructional strategies, student engagement, assessment practices and mastery experiences would directly correlate the results to the TPS implementation expectations. It would be advised to train the same number of teachers in the TPS group as were available in the control group. Larger numbers of trained teachers should create a more accurate picture of the actual efficacy changes that may take place as a result of the TPS training. It would be recommended that future studies separate the control group teachers from the TPS group. This should eliminate confounding effects that casual interaction may have on the end result.

Future studies should also include other forms of assessment instruments to determine student success. It would be beneficial to look at all assessments including: PSSA, 4Sight, classroom assessments and even student grades. Students should be
analyzed on an individual basis and it may be beneficial to focus on those students in the upper and lower 20% of their class. This would allow the researcher to look at student performance outside of the normal expected achievement range and on an individual basis.
References


http://www.pattan.k12.pa.us/teachlead/PennsylvaniaBenchmarkInitiativeBasicInformation


APPENDICES
APPENDIX A: Summer Work Days
Bedford Learning System
Summer Work Days

June 3 days x 3 people, 1 day x 4 people = 13 days
August 1 day x 3 people, 2 days x 4 people = 11 days
Total of 24 days 8:00 to 3:00

1. Math Fact Practice Kits for addition, subtraction, multiplication and division were copied, cut, laminated, and organized.
2. Individual student tracking folders were assembled for Math Fact Practice
3. “Solve the Math Mystery” cards were laminated and organized
4. “Moving on With Tangrams” activities were laminated and organized
5. Eight Versa Tiles activity books, laminated and organized
6. Leveled Problem Solving pages laminated and organized
7. Pentominoes Puzzles laminated and organized
8. Mental Math Problems laminated and organized
10. Buckle Down PA (PSSA Practice) five books laminated and organized
11. Developed and copied tracking forms for all materials
12. Assembled individual student tracking folders
13. Obtained and analyzed individual student needs as determined by May 4Sight test
14. Correlated 2005-2006 PSSA individual student scores with 3rd grade PSSA and 4th grade 4Sight
15. Determined beginning student groupings for instruction
16. Developed Assessment Design for Model Classroom
17. Developed Math Teaching Design for Model Classroom
18. Name stickers for each student printed
19. Student names put into data base for future use
20. Correlated materials to Pennsylvania Assessment Anchors
21. Adapted current math curriculum materials for use with BLS individualized instruction model
22. Developed Help Line Design
APPENDIX B: Model Line Implementation Plan
## Model Line Plan

<table>
<thead>
<tr>
<th>What</th>
<th>Who</th>
<th>By When</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Review/modify/approve implementation plan</td>
<td>Pat/Glenn</td>
<td>3/1/06</td>
<td>Approval</td>
</tr>
<tr>
<td></td>
<td>BLS Team</td>
<td>4/10/06</td>
<td>Concensus</td>
</tr>
<tr>
<td>2) Further training in TPS principles</td>
<td>5th gr. Team</td>
<td>4/10/06</td>
<td>Training Consensus</td>
</tr>
<tr>
<td>3) Review and revise detailed Implementation Plan action steps with</td>
<td>Pat/Glenn</td>
<td>4/28/06</td>
<td>Action Steps Complete</td>
</tr>
<tr>
<td>model line</td>
<td>Madeline, Cathy,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Amy, Ann</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4) Identification of individual student needs in 5th grade math</td>
<td>Cathy</td>
<td>6/06</td>
<td>Info ready to individualize</td>
</tr>
<tr>
<td></td>
<td>5th grade Team</td>
<td></td>
<td>math instruction</td>
</tr>
<tr>
<td>5) Create plan to meet individual student needs in math, including</td>
<td>Pat/Glenn</td>
<td>6/06</td>
<td>Plan of work ready for</td>
</tr>
<tr>
<td>a ‘store’ of assessments, strategies and materials</td>
<td>Madeline, Cathy,</td>
<td></td>
<td>summer curriculum days</td>
</tr>
<tr>
<td></td>
<td>Amy, Ann</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5th grade Team</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6) Adopt/develop ‘store’ of assessments, strategies and materials</td>
<td>Pat/Glenn</td>
<td>8/06</td>
<td>Materials, strategies, pre&amp;</td>
</tr>
<tr>
<td>(includes pre-and post-assessment design)</td>
<td>Madeline, Cathy,</td>
<td></td>
<td>post-assessment design ready for use</td>
</tr>
<tr>
<td></td>
<td>Amy, Ann</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5th grade team</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7) Classrooms and time schedules arranged to facilitate large group,</td>
<td>Madeline, Amy,</td>
<td>School</td>
<td>Classrooms, schedules, store</td>
</tr>
<tr>
<td>small group and individual instruction</td>
<td>5th grade team</td>
<td>Opening</td>
<td>ready</td>
</tr>
<tr>
<td>8) Planned curriculum, pre-/post-assessment cycle and other feedback</td>
<td>Amy</td>
<td>School</td>
<td>System in place and working</td>
</tr>
<tr>
<td>loops begin</td>
<td>5th grade team</td>
<td>Opening-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ongoing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9) Pre-planned instructional strategies (store) applied based on</td>
<td>Amy</td>
<td>Ongoing</td>
<td></td>
</tr>
<tr>
<td>ongoing feedback</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10) When actual student outcome doesn’t match expected outcome, root</td>
<td>Amy, Cathy and ‘help line’,</td>
<td>Ongoing</td>
<td>Problems are identified,</td>
</tr>
<tr>
<td>cause problem solving occurs</td>
<td>others as needed</td>
<td></td>
<td>solved, and do not reoccur</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11) Metric data gathered</td>
<td>Amy, Cathy, 5th grade team,</td>
<td>Ongoing</td>
<td>Data gathered</td>
</tr>
<tr>
<td></td>
<td>4th grade team</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12) Share progress, learning</td>
<td>Amy, Cathy, 5th grade team,</td>
<td>Ongoing</td>
<td>Learning shared</td>
</tr>
<tr>
<td></td>
<td>learning teams, Help Line</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13) Model line evaluation</td>
<td>Amy, Cathy, 5th grade team,</td>
<td>End of</td>
<td>Data reported learning shared,</td>
</tr>
<tr>
<td></td>
<td>BLS team</td>
<td>each</td>
<td>decide on next steps</td>
</tr>
<tr>
<td></td>
<td></td>
<td>semester</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX C: Classroom Model Design
Classroom Design

BLS 5th Grade Math Design-Daily Lesson and Weekly Assessment
(5-15 min.)
↓Group Lesson↓
↓Teacher observed assisted practice↓
(20-30 min)↓↓(Proficient at all other skills)
Prescribed maintenance→→→Independent paper practice (20 min.)
(not proficient at all skills) (Unsuccessful Practice) (Successful Practice)
↓↓↓↓↓↓
Proficient at all other skills

Reteach ←←
(Previous weekly <85%)

Enrichment

-Alternate work
-Refer to tutoring
-Get additional help
-More independent practice
-8:00-8:30 am work w/teacher
-Use ‘teachable’ moments
-Other remediation options
↓↓(5 min.)↓↓
→Individual Math Facts Practice ←

<100% on test←←←←Individual Math Facts Test→→→→100% on test
↓↓↓↓↓↓
Repeat Skills Practice Next in Sequence
↓↓↓↓↓↓↓↓Weekly Practice ←←←←←←
(Weekly)
Less than 85%←←←←←←←←←→Greater than 85%

Identify Problems/Barriers to Success Advance to Next Week
Solve to Root Cause↓↓↓↓
Re-Test↓↓↓↓
Less than 85% ←← Greater than 85%
↓↓Call for Help Line↓
Support (End of Chapter) →→→→Chapter Test←←←←←←←←←
Additional Problem
Solving and Resources
APPENDIX D: Help Line Flowchart
Model Classroom Design for 5th Grade Math
APPENDIX E: Problem Identification Process
The Toyota Way

Problem Identification

How to Identify a Problem

In 5th grade math, it’s a problem when…

- It affects student learning unexpectedly
- A student scores less than 85% on the weekly assessment
- Expected materials and other resources are not available

Remember, in BLS there are 3 general types of problems:

- Work is not done as expected/ customer need is not met
- Work is not done according to the Rules in Use
- Work is not at the Ideal

Help Line Meeting
9/13/2006
APPENDIX F: Problem Documentation Log
# Problem Documentation

## Problem Log

<table>
<thead>
<tr>
<th>#</th>
<th>Date/Time</th>
<th>Prob Id</th>
<th>Problem Description</th>
<th>Date/Time Restored System</th>
<th>Problem Solving Document</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX G: Ohio State Teacher Efficacy Scale
Teacher Efficacy Scale

Teacher Beliefs

Directions: This questionnaire is designed to help us gain a better understanding of the kinds of things that create difficulties for teachers in their school activities. Please indicate your opinion about each of the statements below. Your answers are confidential.

How much can you do?

1. How much can you do to get through to the most difficult students?
2. How much can you do to help your students think critically?
3. How much can you do to control disruptive behavior in the classroom?
4. How much can you do to motivate students who show low interest in schoolwork?
5. To what extent can you make your expectations clear about student behavior?
6. How much can you do to get students to believe they can do well in schoolwork?
7. How well can you respond to difficult questions from your students?
8. How well can you establish routines to keep activities running smoothly?
9. How much can you do to help your student’s value learning?
10. How much can you gauge student comprehension of what you have taught?
11. To what extent can you craft good questions for your students?
12. How much can you do to foster student creativity?
13. How much can you do to get children to follow classroom rules?
14. How much can you do to improve the understanding of a student who is failing?
15. How much can you do to calm a student who is disruptive or noisy?
16. How well can you establish a classroom management system with each group of students?
17. How much can you do to adjust your lessons to the proper level for individual students?
18. How much can you use a variety of assessment strategies?
19. How well can you keep a few problem students from ruining an entire lesson?
20. To what extent can you provide an alternative explanation or example when students are confused?
21. How well can you respond to defiant students?
22. How much can you assist families in helping their children do well in school?

23. How well can you implement alternative strategies in your classroom?

24. How well can you provide appropriate challenges for very capable students?
APPENDIX H: PSSA and 4Sight Comparison
Comparison of 4Sight Predictions to Actual PSSA Scores
Grade 3, 4 and 5  2005-2006 School Year

<table>
<thead>
<tr>
<th>Reading Proficiency District</th>
<th>Math Proficiency District</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prof/Adv</td>
<td>Bas/Below</td>
</tr>
<tr>
<td>4Sight 356</td>
<td>141</td>
</tr>
<tr>
<td>PSSA 334</td>
<td>162</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reading Proficiency BE</th>
<th>Math Proficiency BE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prof/Adv</td>
<td>Bas/Below</td>
</tr>
<tr>
<td>4Sight 281</td>
<td>108</td>
</tr>
<tr>
<td>PSSA 276</td>
<td>113</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reading Proficiency HYLO</th>
<th>Math Proficiency HYLO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prof/Adv</td>
<td>Bas/Below</td>
</tr>
<tr>
<td>4Sight 74</td>
<td>33</td>
</tr>
<tr>
<td>PSSA 58</td>
<td>49</td>
</tr>
</tbody>
</table>
APPENDIX I: Interview Questions TPS Teacher
Interview Questions

**Interview Questions for TPS Teacher**

1. What are the major differences between the TPS model and a traditional classroom model from the teacher’s perspective?

2. Looking back over your experience implementing TPS, how would you describe the impact it had on your professional growth as a teacher?

3. How would you describe the impact that TPS had on student performance, student perceptions, and parent perceptions and on other fifth-grade staff members?

4. Are there any major eye opening experiences that may impact the methods of instruction you choose to practice as a result of this exposure to TPS?

5. Do you have any other significant findings, as a result of your experience with TPS that may be of interest to this study or to the educational research community?