Spring 5-2009

Teacher Evaluation Practices And Student Achievement

Lucy Kay Lyon
Lindenwood University

Running head: Teacher Evaluation and Student Achievement

Teacher Evaluation Practices And Student Achievement

Lucy Kay Lyon

May, 2009

A dissertation submitted to the Education Faculty of Lindenwood University in partial fulfillment of the requirements for the degree of

Doctor of Education

School of Education

DECLARATION OF ORIGINALITY

I do hereby declare and attest to the fact that this is an original study based solely upon my own scholarly work here at Lindenwood University and that I have not submitted it for any other college or degree here or elsewhere.

Full Legal Name: Lucy K. Lyon

Signature: Lucy K. Lyon Date: 8/9/09

TEACHER EVALUATION PRACTICES
AND STUDENT ACHIEVEMENT

Lucy Kay Lyon

This dissertation has been approved as partial fulfillment
of the requirements for the
degree of
Doctor of Education
at Lindenwood University by the School of Education

Dr. Terry Reid, Dissertation Chair	Aug 6, 2009 Date
Dr. Sherry DeVore, Committee Member	Aug. 6, 2009 Date
Dr. Dale Slagle, Committee Member	July 28, 2009 Date

ACKNOWLEDGEMENTS

Thank you to the following people for their understanding and for their support during this endeavor. My appreciation goes to my daughters, Erin and Regina, for understanding the commitment of time and energy to produce this document. Mark has been supportive and provided constant encouragement. My professors have provided guidance and advice. To all the Top Ten schools, congratulations and keep up the great work for our students.

Abstract

Extensive research has been conducted on improving student academic achievement and techniques to improve student learning. There has been little research that addresses the relationship between student achievement and teacher performance. The purpose of this study was to determine the relationship between performance-based teacher evaluation practices and increased student achievement. This study was conducted using the Top Ten Performing School Districts on the Missouri Assessment Program communication arts and mathematics tests and performance-based teacher evaluation systems. A relationship was found to exist between the inclusion of criteria specific to student achievement in the performance-based teacher evaluation program and ranking in the Top Ten on the state assessment. The relationship showed the higher the ranking, the more likelihood of the use of student achievement data in the evaluation process.

TABLE OF CONTENTS

LIST OF TABLES vii

LIST OF FIGURES viii

LIST OF ABBREVIATIONS ix

CHAPTER ONE - INTRODUCTION 1

Background 1

Conceptual Underpinnings 3

Statement of the Problem 5

Purpose of the Study 5

Question 6

Independent Variable 6

Dependent Variables 6

Hypotheses 7

Limitations of the Study 7

Definitions of Terms 8

Summary 13

CHAPTER TWO – REVIEW OF LITERATURE 14

Introduction 14

Accountability on the National Level 14

Accountability on the State Level 19

Quality Education Research 23

Student Achievement 26

Missouri Assessment Program (MAP) 27

Understanding Assessment 31

Teacher Effectiveness and Evaluation 34
History of Teacher Evaluation 54
Summary 59
CHAPTER THREE – DESIGN AND METHODOLOGY 60
Introduction 60
Research Question. 61
Subjects 61
Sampling Procedure 63
Research Setting 63
Research Design Procedure 64
Data Analysis 65
Statistical Treatment of Data 67
Ethical and Political Considerations of the Study 68
Summary 68
CHAPTER FOUR - RESULTS 70
Introduction 70
Results 71
Analysis of Data 76
Research Question 76
Deductive Conclusions 77
Summary 77
CHAPTER FIVE - DISCUSSIONS 79
Introduction 79
Implication for Effective Schools 80

Recommendations....81

Summary....82

References....84

Appendix A: State Statute 178.128 92

Appendix B: Missouri Performance-Based Teacher Evaluation Criteria 93

Appendix C: Survey 100

Appendix D: IRB Approval Form....102

Vita....110

LIST OF TABLES

Table 1: Spearman Rank Order Correlation Coefficient
Raw Data Ranking Used 72

Table 2: Spearman Rank Order Correlation Coefficient
Ranking Data Used 73

Table A1: Timeline of the Study 95

Table A2: Ranking of Buildings by Percentage for Top Ten ... 96

Table A3: Top Ten Buildings Ranked by & Possibility
In Top Ten 98

LIST OF FIGURES

Figure 1. Average Ranking in Top Ten …….74

LIST OF ABBREVIATIONS

APR	Annual Performance Report
AYP	Adequate Yearly Progress
CRT	Criterion Referenced Test
MAP	Missouri Assessment Program
MDESE	Missouri Department of Elementary and Secondary Education
NCLB	No Child Left Behind
SES	Supplemental Educational Services
USDE	United States Department of Education

CHAPTER ONE

Introduction of the Study

Background

The quality of the educational system in the United States has been called into question by legislative leaders for several years. Most recently, the No Child Left Behind Act (NCLB) in 2002 required all students demonstrate proficiency in reading, mathematics, and science by the year 2014. The intent of the No Child Left Behind legislation was to increase the accountability of the Title I programs by requiring states to implement school-wide accountability plans for student achievement (United States Department of Education [USDE], 2004).

Missouri's plan, the Missouri School Improvement Program (MSIP), outlined fourteen areas of accountability with levels of expectation for quality schools. Schools are evaluated annually based upon a set of standards for all districts and an Annual Performance Report (APR) is published yearly. Schools with high levels of achievement, as well as those with areas of deficit, are

identified. A School Report Card is compiled annually through the Missouri Department of Elementary and Secondary Education. This report outlines the areas of student achievement that have been met or not met according to the No Child Left Behind standards. This report card includes the Annual Yearly Progress (AYP) Report on the national standards and the Annual Performance Report that considers the fourteen state standards for accreditation (Missouri Department of Elementary and Secondary Education [MDESE], 2008).

Teacher evaluation in Missouri has been in existence, to some extent, as long as there have been public schools in Missouri. However, it has only been in the last 20 years that there has been some sort of systematic way to evaluate teachers. Missouri Statute *Section 168.128* (see Appendix A) outlined the provisions for teacher evaluation for all public schools in Missouri. The statue stated the board of education for each school district is responsible for maintaining records showing tenure of teaching in the district. The law also reads that each district will conduct a performance-based teacher evaluation. This teacher evaluation must be ongoing and provide specific

information related to the teacher's competency and ability to teach (MDESE, 1999).

This law does not specify the number of times, the specific criteria to use, or how to improve the process. The determination of the implementation of these is left to the discretion of the individual boards of education and even to the discretion of the building principal to interpret the individual criteria and descriptors for each system (MDESE, 1999).

Accountability, on the local school district level, takes into consideration teacher evaluation systems. The systems that school districts use to evaluate both tenured (permanent) and non-tenured (probationary) teachers can affect student achievement. Marzano's research has shown that the quality of the teacher directly affects the performance of students (Marzano, 2003). With the increased level of accountability on the local level, districts are seeking ways to increase student achievement to meet the benchmarks set by No Child Left Behind.

Conceptual underpinnings.

Schmocker's claims that schools today have the opportunity to close the achievement gap and raise student achievement to extraordinary levels quickly are

based on clear evidence (Schmocker, 2006). The evidence is indisputable. Teaching had the greatest impact on student achievement. This sounds simple, but schools today must take a serious look at the instruction that is taking place in the classrooms (Schmocker, 2006). Based upon the research gathered about teaching and learning, student achievement is not making the gains expected. This occurs because instruction is not closely observed or supervised (Schmocker, 2006).

The best explanation Schmocker has for the reason student achievement is not gaining exponentially deals with those who are directly involved not knowing exactly what is going on in the classrooms. Schmocker claims there is a "protective barrier that discourages and even punishes, close, constructive scrutiny of instruction and the supervision of instruction" (Schmocker, 2006, p. 13). This barrier protects those inside of classrooms, as well as insulates the public from knowing what is actually going on inside of the schools.

"This status quo gets enormous help from the machinery that creates the illusion of scrutiny and inspection – namely, teacher and administrative evaluations" (Schmocker, 2006, p. 15). Schools must

challenge the fundamental state of instruction and supervision.

Statement of the Problem

Student achievement is paramount in today's educational society of high-stakes testing. According to the National Assessment of Educational Progress Report of 2002, the achievement gap between white and minority students, as well as between poor and more economically advantaged students, has widened (USDE, 2004). Districts are charged with the responsibilities of meeting the demands of legislation relating to improving student achievement. All educators must be highly qualified in the core academic areas in which they teach. Research shows a clear correlation between the academic achievement of students and the quality of the teacher. Teacher evaluation systems are one way to prove accountability on the local level (Whitehurst, 2002).

Purpose of the Study

The purpose of this study was to determine the effect of performance-based teacher evaluation on student achievement. When school personnel are held accountable for student achievement through an evaluation system, does student achievement increase? Teacher accountability can focus the efforts on actual teaching

performance and provide a structure to improve the teaching abilities of the staff. Effective teaching must be defined, measured, and related to student achievement.

Question

The following overarching research question was addressed:

1. What is the impact, if any, of using student achievement data to evaluate teacher performance during the performance-based teacher evaluation process?

Independent Variable

The independent variable in this study was the performance-based teacher evaluation used in schools that Missouri Assessment Program scores ranked in the Top Ten based on MAP Results for school buildings in the district. Specifically, the instrument was studied to determine the evidence of criteria on the teacher evaluation instrument for that district in the Top Ten ranking.

Dependent Variable

The dependent variable in this study was the student results from the Missouri Assessment Program (MAP). The results from the Missouri Assessment Program areas of mathematics and communication arts were used. The grade

levels for communication arts included grades 3, 4, 5, 6, 7, 8, and 11. The mathematics MAP results were from students in grades 3, 4, 5, 6, 7, 8, and 10. Results from the 2005-06 and 2006-07 school years were used.

Hypotheses

1. There is no effect on student achievement when the performance-based teacher evaluation process specifically identifies and uses criteria related to improving student achievement.

Limitations of the Study

Limitations were considered in this study, ranging from consistent use of the teacher evaluation instrument to interpretation of the criteria in the instrument.

1. This study considered only schools in Missouri that were identified in the Top Ten ranking on the Missouri Assessment Program for mathematics for grades 3, 4, 5, 6, 7, 8, and 10 and communication arts for grades 3, 4, 5, 6, 7, 8, and 10.
2. The years of the study included those schools in the ranking for the 2005-06 and 2006-07 school years.
3. School districts were identified, but individual school buildings were targeted for the survey to gain an accurate interpretation of the district

evaluation instrument at the actual implementation level.

4. Due to lack of reliability and accuracy of the teacher evaluation instrument that a district uses in evaluating its professional staff, this caused limitations within the study. Since there is no state-wide instrument that is to be used by each school district in Missouri, each district may develop its own instrument and conduct the evaluation as it sees fit. This also leads to variations in the interpretation of the criteria in each district, as well as in many buildings in the school district itself.
5. This study does not consider the socio-economic status of the districts in the survey. There are some districts with very limited resources, in terms of administrative staff and opportunities for staff development for understanding teacher evaluation and its implications.

Definitions of Terms

The following terms have been defined to provide for easier comprehension of the study.

Adequate yearly progress (AYP). An individual state's measure of yearly progress toward achieving state

academic standards, as described in the NCLB legislation. AYP is the minimum level of improvement that states, school districts, and schools must achieve each year (MDESE, 2008).

Advanced/Proficient. Two of the four benchmark quartile achievement scores which are calculated by a percent of the raw score on a criterion-referenced test determined by the state as necessary to meet AYP. These are the top two standards of performance for each assessed content area. The other two quartiles are below basic and basic (MDESE, 2008).

Annual performance report. Report submitted by the Department of Elementary and Secondary Education for each public school district in Missouri. The report is based on how well a district has met each of the fourteen standards (MDESE, 2008).

Criterion-referenced tests (CRT). "An assessment that measures a student's mastery of skills or concepts set forth in a list of criteria, typically a set of performance objectives or standards. Such tests are designed to measure how thoroughly a student has learned a particular body of knowledge without regard to how well other students have learned it"(Ravitch, 2007, p. 64).

Formative assessment. Any assessment used by educators to evaluate students' knowledge and understanding of particular content and then to adjust and plan further instructional practices accordingly to improve student achievement in that area (Ravitch, 2007, p. 98).

Missouri assessment program (MAP). One of several educational reforms mandated by the Outstanding Schools Act of 1993. As a result of this act, the State Board of Education directed the Missouri Department of Elementary and Secondary Education to identify the knowledge, skills, and competencies that Missouri students should acquire by the time they complete high school and to evaluate student progress toward those academic standards. The assessment program used is identified as the MAP (MDESE, 2008).

No Child Left Behind Act. A legislative act initiated by the Bush Administration to establish accountability for the nation's public schools through a measurement of Adequate Yearly Progress. Schools and districts are to achieve a goal of 100 percent proficiency in reading, mathematics, and science for every subgroup by the 2013-2014 school year (MDESE, 2008).

Norm-referenced tests. An assessment designed to compare the scores of individuals or groups of individuals with the scores achieved by a representative sample of individual with similar characteristics, members of a so-called reference group. Norm-referenced tests are useful for comparing the performance of students in one school, district, country, state or nation with the performance of students in others (Ravitch, 2007).

Permanent teacher (Tenured teacher). Any teacher who has been employed as a teacher in the same school district for five successive years and who has continued to be employed as a teacher by the school district (MDESE, 1999).

Probationary teacher (Non-tenured teacher). Any teacher who has been employed in the same school district for five successive years or less (MDESE, 1999).

Report card. Under NCLB, states must require districts to publicly report state-mandated assessment information and provide explicit information to students, parents and teachers about the results of student progress (MDESE, 2008).

Rubric (Scoring guide). A set of criteria for evaluating student work or scoring tests is defined as a

rubric. Rubrics describe what work must look like to be considered excellent, satisfactory, or less than satisfactory. In particular, rubrics are needed to minimize subjective judgments of performance (Ravitch, 2007).

School choice. Schools that do not meet Adequate Yearly Progress must inform parents of their right to withdraw their children from the district and place them in a higher performing school without penalty (MDESE, 2008).

School improvement. A term used to designate a Missouri school district or building which does not meet Adequate Yearly Progress for two consecutive years (MDESE, 2008).

Student achievement. This is a definitive measure of a student's academic growth through norm-referenced or criterion-referenced test batteries (MDESE, 2008).

Teacher evaluation. A term used to identify a system of feedback for a teacher that is designed to measure teaching competence (MDESE, 1999).

Teacher evaluation instrument. Instrument and system used to evaluate teachers on a local level (MDESE, 1999).

Top Ten Schools in Missouri. Top Ten Schools in average percent of students scoring at the "Proficient"

and "Advanced" levels on the communication arts and mathematics MAP assessments at grades 3, 4, 5, 6, 7, 8, 10, and 11 (MDESE, 2008).

Summary

This chapter provided an introduction to the study and discussed the methods of data collection. The criteria used during the performance-based teacher evaluation process are used to determine the effect on student achievement. It established the purpose of the study along with the definitions and assumptions. Chapter Two provides an overview of the relevant literature, data, and experiences. The remaining chapters were dedicated to the design, methodology, and analysis of the data.

CHAPTER TWO

Review of Literature

Introduction

As society moves further into a global economy and high-stakes testing becomes an issue in public schools, there is an increased emphasis on comparison of the academic achievement of students on the standardized tests in communication arts, mathematics, and science (MDESE, 1999). Test scores are even used to gauge the value of the future economy of countries based on the quality of their educational programs. It is assumed that countries with students who score the highest are doing a better job educationally and translates to increased competition in the world economy. It is the responsibility of each public school district in the nation to provide a quality education to each student. School districts are seeking ways to improve student achievement (USDE, 2000).

Accountability on the National Level

In the last several decades, the performance of United States' students has fallen in communication arts,

mathematics, and science in comparison with other world economies, according to the National Assessment of Educational Progress (USDE, 2000). As a result, Congress and the President have issued an order for proficiency of students in the areas of communication arts, mathematics and science by the year 2014. This massive bi-partisan legislation, entitled No Child Left Behind, mandated that public schools increase student achievement levels or face sanctions by the federal government (USDE, 2004).

The No Child Left Behind legislation called for students in grades three through eight to test in the areas of communication arts and mathematics. States were allowed to develop and administer their own tests and decide upon the proficiency rating for each subject area and grade level. Missouri used educators, business leaders, politicians, state department specialists, as well as parents in this task. These groups of people met and determined the level of proficiency for each test. Based upon the annual test results, schools are placed on lists as to whether they have met the requirements of No Child Left Behind (USDE, 2004)

The state departments of education, in conjunction with the federal department of education, annually determine if a district has met Adequate Yearly Progress

(AYP). States are mandated to impose sanctions on the school districts that fail to meet the standards set forth by the legislation. There are four factors that are considered in meeting the Adequate Yearly Progress targets. The percent of students meeting the proficiency target is the first factor. Schools must also report that ninety-five percent of the eligible students actually are assessed for each grade level and subject area. Missouri also uses the attendance rate for all students and graduation rate in meeting the requirements for the legislation (USDE, 2004).

In Missouri, the first factor of accountability takes into consideration the Missouri Assessment Program results. The areas of communication arts and mathematics are assessed for all students in grades three through eight. Until the 2008-09 school year, students in grade ten were administered the mathematics assessment while students in grade eleven took the communication arts assessment. Beginning with the 2008-09 school year, Missouri high school students are now responsible for their performance on end-of-course exams for certain courses. For the 2008-09 school years, students who are enrolled in English II, Algebra I, and Biology will take the end-of-course assessments. For the 2009-10 school

year, the courses expand to English I, Algebra II, Geometry, Integrated Mathematics II and III, American History and Government. Student performance on these assessments will determine whether or not a school meets the No Child Left Behind targets (MDESE, 2008).

States were also allowed to determine the incremental percentages for meeting the targets. Since 100% of the students must be proficient in communication arts and mathematics by the year 2014, Missouri phased in the percentages for each year (MDESE, 2008).

Schools that do not meet the annual targets are identified for school or district improvement. Sanctions are imposed on these schools and districts are on a continuum that can be as simple as restricting the use of the federal funds to making major personnel changes or even closing the school and sending students to another school that is meeting the targets. The sanctions for Missouri schools include placing school buildings and/or districts in School Improvement Status (MDESE, 2008). They include:

1. Develop or revise a school improvement plan within three months after identification of status.
2. Notify parents of status with a comparison of the academic achievement with other schools in the

district, reasons for the identification, what the school is doing about the problem, and ways parents can become involved in addressing the academic issues.

3. Offer Public School Choice (PSC) to all students to transfer to another public school within the district.
4. Offer Supplemental Educational Services (SES) to those students eligible. Those who take advantage of SES will have tutoring services paid by the district from their federal funds.
5. Ten percent of the Title I funds must be spent on professional development.(MDESE, 2008)

If the school/district does not meet the Annual Proficiency Target, additional sanctions, including restructuring the school/district, can occur with replacing personnel or even closing a school building or district and allowing the students to attend neighboring districts that meet the annual targets.

The Annual Yearly Progress Report outlines whether or not a district has met the expected progress levels for communication arts and mathematics for its third through eighth grade and high school students. Additional criteria for AYP include disaggregating the student

achievement results to determine if subgroups of students also performed at the expected levels. Subgroups can include English Language Learners, special education, disadvantaged (free/reduced meal), race, etc. These students are expected to perform at the same level as their peers. AYP also considers attendance and graduation rate for public school students. AYP is reported annually for each school district (MDESE, 2008).

Accountability on the State Level

Missouri's accountability system, the Missouri School Improvement Program, falls in line with the requirements of the No Child Left Behind Legislation. This program has the responsibility of reviewing and accrediting the 524 four school districts in the state of Missouri within a five-year cycle. It is designed to promote excellence in the public schools in Missouri. The Missouri School Improvement Program has the dual responsibility of ensuring all public schools meet certain minimum standards and strive to achieve excellence in an increasingly competitive world. The goal of the MSIP process is to guide schools in their school improvement efforts. There are three sections to the standards for MSIP: Resource, Process, and Performance (MDESE, 2008).

The Resource Standards address the basic requirements that all districts must meet. They are quantitative in nature. These include program of studies for students, class size and assignments, professional support staff, administrative staff, and certification and planning time. Missouri districts report information to the Department of Elementary and Secondary Education bi-monthly through the Core Data collection system. The resource standards outline the course offerings for elementary, junior high, and high school students. Class size and enrollment data include a minimum standard and a desired standard for the number of students enrolled in K-12 classes. All schools must meet the minimum standard but strive to meet the desired state standard. Professional support staff delineates the librarian/media specialist and guidance and counseling staff in student/teacher ratios for these areas. The number of administrative staff in the central office is determined by the number of professional staff members in the district. The ratio of students to number of principals is kept to a consistent ratio. All professional staff members (teachers) must have the appropriate certification to teach as well as have a minimum of 250 minutes a week of scheduled planning time. All this data

is collected through the Core Data collection system (MDESE, 2008).

The Process standards delineate instructional design and practices. The Process standards deal with curriculum and assessment. Instructional programs, resources, climate, and orderly and safe schools are also scrutinized. Data is gathered through surveys or interviews. Professional development and teacher training is taken into consideration in this area. Differentiated instruction, taking into consideration the disabled, gifted/talented, career and preschool students, is in this area. Parent and community involvement is studied. Additional school services (nursing, transportation, board of education, facilities, and food service) must meet the standards put forth by DESE.

In Standard 6.5, "The district has created a positive climate for learning and established a focus on academic achievement. 2. Teachers and administrators are accountable for promoting student success and reducing student failure" (MDESE 2008), it became apparent that school districts hold teachers and administrators accountable for student achievement. In the Fourth-Cycle Report Writing Form (2009), districts are to provide evidence that student performance data is used in the

teacher and principal evaluation process. Districts are to provide the evaluation team with examples of teacher and principal evaluations with specific criteria related to student achievement. (MDESE, 2008)

The Performance Standards of the MSIP process have taken a much larger role than the Resource and Process standards in Missouri's accountability system. Accreditation is determined by the performance level of the students in a school district in Missouri. If a school fails to meet the standards, it will become unaccredited with sanctions as dire as consolidation with another school district (MDESE, 2007a). The performance standards use five years of data to determine whether a school district has met that standard. These are published by DESE in the Annual Performance Report.

This accountability system takes into account student achievement at all levels in communication arts, mathematics, and science. Other factors considered, just as in the federal legislation, are graduation rates and attendance rates of students in the public schools. Public school systems in Missouri face these accountability measures on an annual basis in the fall of each school year. The local media compare the findings of

the local districts, touting the successes or failures of the public schools.

This accountability has trickled down to the local levels with Boards of Education holding superintendents and building administrations responsible for gains or losses in student achievement. This is further brought down to the level of the classroom teacher. Many researchers have stated that student performance can be directly tied to teacher performance and that in order to see true improvement in student achievement the classroom teacher must be held accountable (Toch and Rothman, 2008).

Quality Education Research

Research during the 1970s and 1980s brought forth the importance of the classroom teacher in student academic achievement. Brophy and Good (1970) suggested that teachers may differentiate their behaviors toward students based on their expectations. They suggested that students will perform to the expected levels of the teachers, given the right conditions. Benjamin Bloom (1981) began the mastery teaching movement which proposed that ninety-five percent of the students can learn any subject to a high degree given sufficient time and appropriate instruction.

In the 1980s the educational reform movement began with the Effective Schools Movement. Edmonds, Lezotte, and others identified the ingredients of an effective school. Teachers spent more time teaching and students spent more time learning in these schools that were identified as effective. There was also maximum teacher/student interaction with focused lessons tailored to individual student needs. Levine studied effective schools and determined high expectations, frequent monitoring of student learning, and frequent evaluation were essential to increased student achievement (Marzano 2003).

In 1983, the report *A Nation at Risk: The Imperative for Education Reform* addressed the concerns of the nation on falling standards in the public schools. The report outlined recommendations for improving the public schools, one being to improve the teacher training and preparation. This report held schools and teachers accountable for student learning.

In the 1990s, Richard DuFour and Robert Eaker began the Professional Learning Communities movement. Teacher collaboration, high expectations, and clear goals for student achievement are all part of this process to improve academic achievement of students. The basic

premise of the professional learning community comes from the business world regarding how organizations learn. This study of good business practice and what makes a difference has been applied to the educational realm. Professional learning communities are based upon two assumptions. First, knowledge is learned and must be shared through critical reflection. Secondly, actively engaging teachers in the collaborative work will result in better student achievement (Vescio, Ross, Adams, 2006).

Reflective dialogue is essential to a professional learning community. Educators must work together to answer clarifying questions such as the following:

- What do we want all students to learn?
- How will we know when each student has mastered the essential learning?
- How will we respond when a student experiences initial difficulty in learning?
- How will we deepen the learning for students who have already mastered essential skills and knowledge? (DuFour, 2005)

Once this process began, educators realized the importance of time and support during the school day. Staff needed time to work collaboratively and learn the

best practices to support viable instruction. Teachers needed time to reflect upon the student learning. They were also responsible for sharing solid instructional practices with their colleagues.

Assessment became a large part of the professional learning community. According to Rick Stiggins (1995), the reasons for assessment are two-fold: to gather evidence of student learning and to motivate student learning. Professional learning communities help create a culture of assessment for learning instead of the traditional assessment of learning. Educators shift from the summative assessments to more productive formative assessments. Effective use of classroom assessments lead to clear and appropriate learning targets for students, increased accuracy of assessments, continuous feedback, and more student involvement (DuFour, 2005).

Student Achievement

Student achievement and its measure have changed significantly since the passage of the No Child Left Behind Legislation in 2002. Because of the amount of information available on how students learn, students must be taught and assessed on how to think, reason, and apply learning, not just the simple memorization of facts. Missouri's assessment system, the Missouri

Assessment Program, began during the 1990s. This development was in response to the Outstanding Schools Act of 1993 which directed the Missouri Department of Elementary and Secondary Education to identify knowledge, skills, and competencies that Missouri students should acquire by the time they complete high school. DESE was also given the task of developing an assessment program that outlined student progress toward those academic standards. The Show-Me Standards were then developed. These were further broken down to the Curriculum Frameworks to provide guidance to districts in planning the curricula designed to ensure students were progressing to meet the standards.

Missouri Assessment Program (MAP)

The No Child Left Behind Act of 2001 required all states to annually assess student learning in communication arts and mathematics at grades three through eight and at a high school grade by the end of 2005. In preparation for these assessments, the Missouri Department of Elementary and Secondary Education delineated the Curriculum Frameworks to address these assessments to provide guidance for the teachers. Grade-level expectations outlined the specific course and grade-level objectives were designed to align with the

incoming assessments. Missouri is currently in the second revision of the grade level expectations to align with end-of-course exams for high school students (MDESE, 2008).

Initially, the Missouri Assessment Program (MAP) was designed to assess students in the areas of communication arts, mathematics, science, social studies, fine arts, and physical education in bench-mark years. As the NCLB standards were put forth, social studies, fine arts, and physical education have been dropped from the program for Missouri students. In recent action by the Missouri State Board of Education, high school students will no longer be administered the MAP tests at grades ten and eleven but will move to an end-of-course exam at any grade level upon completion of the specified courses of communication arts, mathematics, and science. NCLB standards also require all students in all grades three through eight be assessed in communication arts and mathematics and two benchmark years for science. Missouri opted to assess students in grades five and eight for science in the spring of 2008 (MDESE, 2008).

The Missouri Assessment Program (MAP) assessments are comprised of three types of items: 1) selected-response, 2) constructed response, and 3) performance

events. Selected response items are multiple choice which present students with a question followed by four or five response options. These questions are nationally-normed through the McGraw-Hill Terra Nova Assessment. The constructed response items require students to supply (rather than select) an appropriate response. Students might be asked to supply a one-word answer, a sentence, or show their work in solving a problem. The performance events measure students' knowledge and their abilities to apply the knowledge in problem situations. Most of these are multi-step problems requiring a higher level of understanding. While there is an understanding that certain facts must be understood by all students, application and problem-solving are addressed in the assessment program (MDESE, 2008).

These assessments are scored by the CTB McGraw-Hill Company as well as Missouri teachers and professional scorers. The selected response items are scored by CTB McGraw-Hill and reported in percentiles, comparing the student to those in the norm group. The constructed response and performance event items are scored by the professional scorers and Missouri teachers. Teachers spend two weeks during the summer months training to

score and then actually scoring items for the state department of education (MDESE, 2008).

Two types of scores are reported to indicate a student's achievement on the Missouri Assessment Program (MAP) test: (1) a scale score and (2) its associated level of achievement (MDESE, 2007a). A scale score indicates a student's total performance for each content area assessed by MAP. A higher scaled score indicates higher levels of achievement while a lower score indicates the opposite. There are four levels of achievement on the MAP: Basic, Below Basic, Proficient, and Advanced. Cut benchmark scores for each level were identified by Missouri citizens and teachers and reflect the expectations of each group of what the students at each level should know and be able to do. Studies indicate the MAP test is closely aligned with the National Assessment of Educational Progress (NAEP) test. Missouri has conducted extensive studies on the reliability and validity of the MAP test with annual technical reports published in conjunction with CTB McGraw-Hill (MDESE, 2007a).

NCLB mandates that all students test proficient by the year 2014. Missouri has designated a level of

proficient or advanced on the MAP to determine the percent of students at the NCLB proficient level.

Annually, the Missouri Department of Elementary and Secondary Education submit a comprehensive report of the "Top Ten Lists Based on MAP Test Results" in Missouri. These schools are listed in order of the percent of students in the Advanced and Proficient levels on the communication arts and mathematics MAP assessments. The schools are separated based upon the number of students in each building. The breakdown categories are 1) less than 250 students, 2) 250 to 500 students, and 3) over 500 students (MDESE, 2007a).

Understanding Assessment

Teachers must understand the purpose of assessment as well as the instrument that is used to obtain the information or skills. There are norm-referenced achievement tests, criterion-referenced tests, and other types of student assessment. Norm-referenced tests are used in schools to provide information on how well the student compares to other students in the same grade level across the country. According to Tucker and Stronge, norm-reference tests usually answer the following questions related to student learning:

1. Where does a student stand in a given area of achievement in relation to other students and compared to other students and compared to the norm group of students?
2. How does the overall achievement in one teacher's class compare with that of another's?
3. How does the achievement in the given content area for students in the selected school district compare with the national norms or with another school district? (Tucker & Stronge, 2005, p. 18)

The Terra Nova portion of the Missouri Assessment Program is a norm-referenced test. A percentile score reflects how the students perform in relation to a control sample of other students in the nation (MDESE, 2008).

Criterion-referenced tests are also used on the MAP. Criterion-referenced tests measure the student performance to indicate how much has been mastered by the student. These tests are designed to determine whether students have reached an established level of learning in an area. This is the constructed response and performance event portions of the MAP (MDESE, 2008). Again, Tucker and Stronge explain the questions criterion-referenced tests answer:

1. What is the student's level of knowledge in the domain (e.g., what percentage of the problems of a given type can we expect the student to solve correctly?)
2. What are the specific strengths and weaknesses of a given school program or curriculum?
3. What specific changes in student performance have occurred as a result of changing the curriculum or program?

Teachers may also develop local assessments to measure student learning. Other assessments that teachers can use include writing samples, student portfolios, and other performance-based assessments. Teachers must understand each type of assessment in order to gain the information needed for student learning.

Research over the past thirty years has shown there is a correlation between teacher performance and student achievement. These improvement efforts focused on teacher preparation, staff development, and pedagogy. Little emphasis has been placed on teacher evaluation in the past.

According to Holland and Adams (2002), evaluation has traditionally placed teachers in a relatively passive role. Many evaluation systems rely on annual observations

and outdated checklists with no alignment to the teaching standards expected in classrooms. Administrators conducted an observation, wrote a review of the observation, and conducted pre- and post- observation conferences to provide feedback to the teacher. Tucker and Stronge (2005) concluded that the observations were conducted too seldom to provide suggestions that could be tried and reevaluated by the teacher and administrator.

Teacher Effectiveness and Evaluation

Ineffective teacher evaluation systems are more costly than effective ones, according to Danielson and McGreal. Poor evaluation systems neither improve the instructional skills of teachers nor permit the dismissal of ineffective ones. When examining current practices and determining the success of teacher evaluations systems, Danielson and McGreal point out a clear sense of purpose should govern the design of teacher evaluation systems. A teacher evaluation system should screen out unqualified people from certification and the selection process. It should also serve to recognize and reinforce outstanding service.

Although some estimate that incompetent teachers only constitute 2-3% of the teaching population, their presence tarnishes the reputation of the entire

profession. They fail to serve the students and cause parental dissatisfaction with the public schools. Reasons for incompetence are as varied as the number of teachers who exhibit these traits. However, they can fall into three general categories which include influences from non-job related factors, failure of the supervisor to provide assistance, and personal shortcomings of the teacher. Administrators are obligated to confront poor teaching performance and to provide assistance with the deficit. Poor performance can be a result of lack of preparation, deficiency in teaching skills, inability or lack of knowledge of how to control student behavior, poor judgment, and excessive absences (Sawa, 1995).

Administrators, typically, use four different measures to determine the effectiveness of teachers. Supervisory observation is one method that is used to identify incompetent teachers. These can be both formal (scheduled) and informal (unscheduled). Complaints from parents or students can also be an indicator of teacher incompetence. The administrator must weigh the complaints to determine if there is a source of contention between the teacher and person who files the complaint. Complaints from colleagues can provide insight for teacher competence. Student test results, longitudinally

studied, can be a means to identify incompetence (Sawa, 1995).

Several studies over the past twenty years have come to the conclusion the purposes of teacher evaluation systems are varied. Bolton, Denham, Harris and Redfern agree that the major purpose of evaluation is to have a process to provide opportunity for supervisors and teachers to work together to enhance and improve classroom instructional practices. This process will also allow for assistance to those marginal teachers in a structured, systematic way (Stronge 2007).

According to Sawa (1995), there are recommended steps for confronting a teacher with accusations of incompetence. The first step is to gather information; this can be done by talking to others who can be colleagues, parents, or students. The administrator must organize the information and, unfortunately, wait for a specific incident. At that time, a meeting is scheduled with a follow-up letter sent to the teacher outlining the discussion. The administrator must next monitor the situation, developing a file in which specific steps are outlined for improving the situation. A major responsibility on the part of the supervisor it to

continue monitoring the situation, detailing whether or not the situation improves to the level of expectation.

The next step in this process is to determine who the district will retain, transfer or dismiss. With a variety of compensation packages available for educators, the teacher evaluation system may provide guidance on making informed judgments to allow teachers to be eligible for merit pay plans or career ladder plans (Sawa, 1995).

The last function of the teacher evaluation system, according to these researchers, is to provide information to determine the extent of the implementation of the professional development of the district. During the evaluation cycle, it becomes apparent which teachers have used the acquired skills and knowledge that have been presented during the professional development activities (Stanley & Popham, 1988).

Ellett and Garland's study in 1987 surveyed superintendents from the 100 largest schools and compiled data on their teacher evaluation systems. Analysis of the data indicated that most teacher evaluations emphasized the summative rather than the formative purpose of evaluation. The evaluations were used to determine employment status rather than how to improve teaching

through use of professional development. Most of the systems did not include requirements for establishing performance standards or include training for those who were conducting the evaluations as well as those who were to be evaluated. External or peer evaluation was virtually non-existent, while superintendents found their evaluation systems to be more favorable than the researchers compiling the data (Mathers & Oliva, 2008).

Ten years later in 1997, a follow-up study was conducted to Ellett and Garland's findings. The survey was adapted to measure superintendents' opinions about the effectiveness of the teacher evaluation system. The findings included not much had changed from the earlier study in terms of the components and reasoning for the evaluation, but the viewpoints of the superintendents had. They were not satisfied with the status quo. They wanted to revisit the evaluation system and process and revise the tools and procedures (Mathers & Oliva, 2008). Superintendents were recognizing a need for further data from the evaluation systems, but no major changes had been made in the previous ten years to address these concerns.

In 2007, a major study released by the McREL Midwest collected teacher evaluation policies from six states –

Illinois, Indiana, Iowa, Michigan, Minnesota, and Wisconsin. A representative sample of teacher evaluation policies was studied from each of the seven states. During the study, they found that the administrators, usually principals and vice-principals, were responsible for conducting the evaluations of the staff, but fewer than 10% of the policies required evaluator training. These evaluation cycles usually differentiated the evaluation frequency based upon teacher experience, whether the teacher was considered a tenured (permanent) teacher or a probationary teacher. The timelines for evaluation were also listed in about one-half of the policies studied. Administrators were to evaluate in the fall and the spring with summative evaluations conducted in the spring, usually with classroom observations, both scheduled and unannounced.

McREL also reported that over half of the policies identified the type of instrument to be used; however, most of those used the same instrument for all staff, regardless of a teacher's years of experience or subject area. Less than one-third of the policies identified the procedures of how to communicate the evaluation process and procedures to teachers. The most common methods were including the policy in the teachers' handbook,

mentioning it at a group or one-on-one faculty meeting, and even writing it on contracts. The results were similar when an examination was conducted of the common practices of how to share evaluation information with staff members. Most simply have the requirement that both the teacher and administrator will sign off on the summative evaluation after review.

The four most quoted ways of administrators' use of the evaluation information included using the evaluation to drive personnel decisions, suggest improvement for the teacher, set professional development goals, and determine remediation and follow-up procedures for teachers with unsatisfactory evaluations. Little guidance was found to evaluate the specific teacher behaviors and characteristics. Other areas to consider were content and pedagogical knowledge, classroom management skills, effective lesson planning, and fulfilling professional responsibilities. One half of the policies included how well teachers adjust instruction based upon student assessment results (Mathers & Oliva, 2008).

Evaluations should not only provide guidance on how to become a better teacher, but also commend good work with students by the teachers. Based upon the results, teacher evaluation can also drive professional

development for the staff. According to Toch and Rothman (2008), teacher evaluations are the powerful levers that administrators should use to improve the quality of teaching in the classrooms. Administrators should use this lever for school improvement by targeting specific areas for each staff. Areas of concern can be addressed through training and mentoring.

Stanley and Popham note that during the last twenty years in education, there has been a clear move from the volunteering aspect of staff development to the requirement of all staff to participate. This not only includes the teachers, but the administrators as well. This required participation builds consistency between and among the different buildings and organizational structures in a school district. This clear level of accountability for the professional development adds to the administrator's ability to help address areas of concern with the marginal teacher, while providing a common understanding of valid, research-based practices to enhance student achievement (Stanley & Popham, 1988).

Marx (2007) went further in developing guidelines for principals to work with an effective evaluation system. He suggested, in the initial phases of developing an evaluation instrument, one must start with a common

framework or model that identifies good teaching. In Marx's view, involving the teachers in this process to ensure understanding of the model was important. Teachers will also accept the framework more readily if they have had a part in its development.

Successful evaluations and supervision depend upon the quality of the evaluation instrument and the method of gathering the data for the instrument. If the criteria are clear and understood by both the teacher and supervisor, the evaluation instrument will be more accurate and meaningful. Clear, visible, and appropriate criteria for the function of the instrument are essential to the success of the process (Stanley & Popham, 1988).

Marx (2007) recommended that educators determine the purpose of the instrument. Will it be to improve practice which involves formative evaluations? Will it guide in making decisions about retention, advancement or dismissal of teachers which will involve summative evaluation? If the purpose of the evaluation and the method of addressing the accountability are understood by both the teachers and administration, these methods and procedures will be in place to allow for a successful effective tool to develop (Marx, 2007).

Boyd (1989) believed teacher evaluation systems should serve two purposes: to measure teacher competence and to foster professional development and growth. In Boyd's research, he concluded teacher evaluation systems should give teachers useful feedback on classroom needs, provide opportunity for learning new techniques either with the principal or other teachers, and provide support for making these changes in the classroom (Boyd, 1989).

In Toch and Rothman's (2008) research of the nonprofit National Council on Teacher Equality (NCTQ) report of 2007 which studied the nation's fifty largest districts, they found that most union contracts dictate the professional requirements for teachers. It also found that only two-thirds of the contracts required teachers to be evaluated at least once a year. One-fourth of the teachers in this study required evaluations only once every three years (Toch & Rothman, 2008).

The Toch and Rothman study (2008) also showed the evaluations themselves were of little specific value because they did not focus on the quality of teacher instruction. These evaluation instruments were more of a checklist of classroom conditions and teacher behaviors. With these conditions, it was easy for teachers to earn

high marks regardless of whether or not their students were learning.

Using multiple and variable sources of data will enhance the quality of the teacher evaluation instrument. Marx (2007) concluded evaluation instruments in which the principal makes one to two classroom visits using a rating form or anecdotal record is inaccurate and unreliable. He suggested walk-through techniques that can produce more reliable and useful data because they sample classroom behavior more reliably over time. This method is also less intrusive during ongoing instruction (Marx, 2007).

Procedures used to gather data can provide a more accurate view of teacher quality. The most common form of data collection is observation of classroom activities. The goal is to obtain a representative sample of a teacher's performance in the classroom. This must be done, according to Boyd, with multiple opportunities and in a consistent method. Principals may also review lesson plans and classroom records to gain information on the effectiveness of teachers. Lesson plans reflect how well a teacher has thought through the instructional goal. Looking at tests and assignments will give the evaluator

an idea of how the teacher has linked lesson plans, instruction, and assessment (Boyd, 1989).

Common evaluation tools can include lesson plans, self-assessments, portfolio assessments, classroom observations, student achievement data, and student work samples. According to Mathers and Oliva (2008), lesson plans give insight to a teacher's ability to prepare to deliver content, scaffold student skills, and manage the classroom environment. Districts can use rubrics to evaluate lesson plans. Most districts, however, do not require lesson plans to be used as a part of the teacher's evaluation (Brandt et. al, 2007).

The level of planning that a teacher uses to drive instruction is one aspect of good teaching. Effective lesson plans link the student learning objective with the teaching activity. There must be a connection with prior student learning to the taught application or skill in the lesson's objective. The objective must have a strong correlation to the district and state standards. Lesson plans can also describe the teaching practices to maintain student interest and attention. This will help diminish potential classroom management problems. Lastly, lesson plans can provide guidance on how to differentiate the instruction for students with special needs. The

evaluator must remember, however, the lesson plan is just that – a plan for instruction. It must only be scored based on the rubric (Stronge, 2007).

Classroom observation allows evaluators to link lesson plans with actual practice. The classroom observation is the most commonly used tool for evaluating teachers. Evaluators can capture information about teachers' instructional practices with classroom observation. The limitations of this evaluation tool include poorly trained observers and brief, inconsistent observations that create biased results (Shannon 1991).

Several researchers have concluded that student achievement is related to teacher competence in teaching. Wittrock (1986) found that student achievement is tied to the teacher and he/she has a definite impact on student expectations and school ability. The research found that students achieve more when systematic teaching procedures are used. When small increments are applied following each step, this led to greater achievement gains.

Research has repeatedly proved effective teachers have more orderly classrooms. There are more on-task behaviors in those classrooms. The classroom environment must be conducive to learning with neither too much criticism nor praise. More effective teachers have high

levels of student engagement, cooperation, and success. Instructional strategies with high expectations and high content serve the teachers and students in the learning environment.

Teacher expectations are reflected in student work. High achievers were, often times, given specific and sincere praise. The opposite was true of low achievers. High achievers were receiving more frequent and informative feedback, more attention, and treated with more respect. Wait time is longer for low-achieving students.

Effective instructional strategies are essential to student success. Training is necessary for the staff to make the changes in curriculum and strategies. Teaching is prescriptive in its methods and expectations.

Observations are usually conducted by the administrator. Teachers have a high regard for evaluators who possess a deep knowledge of curriculum, content and instruction. These evaluators must also be willing to provide suggestions for improvement. Researchers suggest that multiple evaluators are an alternative to the administrator as the sole evaluator. These multiple evaluators can be peers who have an instructional background, content knowledge, and teaching experience

similar to the teacher being evaluated. These observations would provide specific data for the teacher on the instructional practices being used in the classroom (Mathes & Oliva, 2008).

The length of time, as well as the number of observations conducted, lends themselves to gain a more accurate picture of what is happening in classrooms. Research from Denner, Miller, Newsome, and Birdsong in 2002 suggested that when observations occur more frequently and are longer, their reliability and validity improves.

Non-tenured teachers are normally evaluated annually, while tenured may be on a three to five year cycle for evaluations. This ultimately is not the best way to measure teacher performance if the evaluation captures only one moment in time and the instrument is weak in its interpretation. Both should receive frequent evaluations, according to Mathes and Oliva (2008), as many as five times annually.

Administrators use the teacher evaluation instrument to gather data for both formative and summative evaluations for the staff. According to Popham (2008), formative and summative evaluations focus on different tasks. Formative evaluations, like formative assessment

in the classroom, focus on improvement of skills. These are done more often. For the principal, the focus is on improving instructional skills. The summative evaluation deals with more of a final, summative assessment that draws on formative data. It is not primarily improvement-oriented since it deals with more of a final decision related to evaluation. Principals must understand the distinct difference in each type of data collection, as the teacher must know the difference when administering assessment to the students and the reasons for each type of assessment (Stanley & Popham. 1988).

Self-assessment is another evaluation tool administrators can use to effectively evaluate teachers. Reflection is the process in which teachers analyze their own instruction. This can be accomplished through professional conversations with other teachers during grade-level or subject-area meetings or even through pre-observation and post-observation conferences with the evaluator. Portfolio development can be used by teachers to determine their effectiveness. This can lead to personal professional development plans in which the teacher and evaluator outline a plan of improvement in instructional practices for the teacher (Brandt, et. al, 2007).

Personal reflection has its strengths and drawbacks. The reflection may encourage teachers to continue to learn and grow throughout their careers. This can be done with video-taping classroom lessons and reviewing them, either with colleagues or alone. This can also be used in conjunction with data collection in which the teacher and evaluator reflect about the behaviors and practices observed. This practice does require time and, more importantly, a cultural norm that will support it. The trust factor weighs more heavily on this practice (Mathers & Oliva, 2008).

Portfolio assessment can consist of several types of teacher classroom performance, such as lesson plans, videos of lessons taught, reflection and self analysis of teaching practices, examples of student work, and examples of teacher feedback given to students. This practice enables teachers to reflect on their own instructional practices, enabling them to identify instructional strengths and weaknesses. Focused professional development can be planned from this self-refection. Portfolios are useful, according to Danielson (1996), because they allow administrators/evaluators to review non-classroom aspects of instruction as well as provide teachers the opportunity to reflect on teaching

practices, using the documents contained in the portfolio. Teachers become active participants in the evaluation process when portfolios are used (Danielson, 1996).

Research on the use of portfolios has no conclusive findings. The reliability of this method has not been consistently established, not even when the use of portfolios actually reflects what is going on in the classrooms. No conclusive evidence exists that the process of developing a portfolio and being evaluated by that system leads to improvement in teaching practices and student learning (Attinello et. al., 2006).

The least-used method of collecting data for teacher evaluation is the use of student achievement data. To help determine the effect of teaching on student achievement, some systems use a statistical technique to analyze the changes in standardized scores from one year to the next. The proficiency standard can be used as well as the growth model that measures changes in student performance over a period of time. The use of standardized student test scores enables schools to measure the impact of teaching on student achievement. This builds on the investment in student testing. These items on standardized assessments have been tested for

issues of fairness and appropriateness. This lends to consistent data. The evaluator can determine the relationship between student achievement gains, teachers, and schools (Braun, 2005).

The difficulty of using standardized assessment data for teacher evaluation lies in the instrument used. These tests only measure a portion of the curriculum and teachers' effect on learning. It is difficult to differentiate the elements of learning that affect student achievement or determine which have a positive impact on student achievement. An additional concern is that not all teachers can be assessed using standardized student achievement. Not all grade levels and subject areas are tested annually (Mathers & Oliva, 2008).

An alternate method to determine the effectiveness of teaching practices on student achievement is the use of student work samples. This method provides a more insightful review of student learning over a period of time. Use of this specific data can determine which elements of teaching relate more directly to increasing student achievement than just the standardized scores. This, however, can be time-consuming with the issues of validity and reliability coming into question with work samples as opposed to standardized test results. The use

of well-developed rubrics that clearly outline the criteria for rating student work samples can provide consistent data on student achievement (Mathers & Oliva, 2008).

Leadership responsibilities play an important role in the effectiveness of the teacher evaluation system. Leaders must have a strong, positive role in the evaluation process. The principal must be able to collaborate with teachers and provide useful feedback. Studies focused on teacher perceptions of evaluation found, according to Marx (2007), effective feedback was the most important contributor to changing teacher behavior. This can be challenging at the secondary level with subject area expertise coming into play.

An evaluation instrument must be reliable and valid. It is considered reliable if two or more evaluators use the same instrument and come to the same conclusions. There must be clearly identified criteria that are as objective as possible that require little interpretation. This is accomplished by carefully developing the instrument and training the observers. Validity of the evaluation instrument rests with the interpretation of the criteria. The instrument is valid if it measures what it says it is to measure (Mathes & Oliva, 2008).

Extensive research in the area of improving teacher evaluation systems has been conducted in the last ten years. In the research brief from Marx (2007), it was noted Bradshaw and Joyner (2002) had done extensive research of literature identifying criteria that can be used to develop a more effective teacher evaluation system. They suggest linking evaluation to school goals, gathering and using data on student performance, establishing feedback mechanisms, and including ways to meaningfully involve teachers in the process (Marx, 2007).

All teacher evaluation systems must also be able to withstand professional scrutiny and stand up in a court of law. If the evaluation is to provide evidence for termination of incompetent or unproductive personnel, it must be able to stand in a court of law.

And finally, the evaluation should unify teachers and administrators in their collective efforts to educate students. The goals for both the administrator and the teacher should be to increase student achievement.

History of Teacher Evaluation

Teacher evaluation in Missouri has been in existence, to some extent, as long as there have been public schools in Missouri. In the last 20 years in

Missouri, there has been significant progress in the systematic way to evaluate teachers.

According to Danielson and McGreal (2000), during the 1940s and 1950s, educators emphasized the traits teachers naturally possessed to determine the effectiveness of the teacher. These traits include voice, appearance, emotional stability, warmth, truthfulness, and enthusiasm. Educators believed that those who possessed these traits were likely to perform more effectively with students. There was no real evidence to link these variables to good teaching or to improve student learning (Danielson & McGreal, 2000).

A revolutionary movement of educational practices occurred during the 1960s and 1970s. Researchers began taking a serious look at teaching and student learning. Clinical evaluation processes were developed with observation and rating instruments used to determine what was occurring in classrooms (Danielson & McGreal, 2000).

In 1987, the National Board for Professional Teaching Standards developed a performance-based evaluation system to recognize advanced competence among experienced teachers. This came about through discussions with administrators and teachers on more meaningful standards. These standards were used in thirty-three

states. These standards were moving from the traits teachers should possess to more in-depth views of teaching and learning. Teachers' commitment to students and learning, knowledge of subject matter and how to teach it, managing student behavior, and learning from experience were all standards recommended from the National Board (Weiss & Weiss, 1998).

Over the past twenty-five years, different teacher organizations and research groups have created core standards for beginning teachers and experienced teachers. This delineation in expectations evolved through research and practical application of teacher evaluation systems (Weiss & Weiss, 1998).

Updated in 1983, Missouri Statute Section 168.128 outlines the rules and regulations regarding teacher evaluation for the public school teachers in Missouri. The length of employment is included in the records that must be retained on each employee. The local Board of Education is also responsible for developing a performance-based teacher evaluation. The only criteria related to this statue include the ongoing nature of the evaluation system. The district must determine the standards, frequency, and interpretation of the standards (MDESE, 1999). This law outlined the requirements of all

Missouri public school districts to provide teacher evaluation for all staff members. It did not specify the number of times, the specific criteria to use, or how to improve the process. It did give the State Department of Elementary and Secondary Education the responsibility to provide suggested procedures for teacher evaluation. The specifics of the evaluation system and instrument were left to the discretion of the individual boards of education (MDESE, 1999).

Missouri's teacher evaluation program, recommended by the Department of Elementary and Secondary Education, identified the criteria districts should use to evaluate professional staff. Mandated by legislation in 1993, the performance-based teacher evaluation process was developed as a guide for school districts. Finalized in 1998, the *Guidelines for Performance-Based Teacher Evaluation* provided twenty standards and criteria used in evaluating a teacher's performance. The standards are related to the teacher behaviors. The criteria for teacher evaluation was rated based upon administrator observation and documentations. The standards and criteria for performance-based teacher evaluation identified the actions of the teacher.

School districts in Missouri have the flexibility to use this recommended system or develop one of their own. This system did not directly tie increases in student achievement to teacher evaluation. It outlined criteria for effective teaching practices that are linked to increasing student achievement. Missouri Revised Statute, Chapter 168.128 (1983) mandated a comprehensive, performance-based evaluation for each teacher in the school district. The only stipulations of this statute were that the evaluation must be on-going and specific to demonstrate standards of competency and academic ability. There was no guidance as to the specific criteria or how to interpret the criteria for consistency.

Missouri statute also defines the status of teachers in public school systems. A "permanent" teacher is any teacher who has been employed and who is thereafter employed in the same school district for five successive years (Missouri Statute 168.104). This tenured teacher has an indefinite contract with the school district. A probationary teacher is any teacher who has been employed in the same successive school district five years or less. The probationary teacher must receive notification of re-employment annually until tenure is attained. *The Guidelines for Performance-Based Teacher Evaluation*

delineate the increments of the teacher evaluation cycle for tenured teachers as every five years and probationary teachers on an annual basis.

Summary

This chapter involved a thorough examination of the No Child Left Behind (NCLB) mandate and its effect on public school accountability. Teacher evaluation systems and instrument recommendations were studied to provide guidance for school districts. The historical aspects of teacher evaluation were outlined to get a perspective on the changes in criteria for the evaluation instrument and the role of the teacher and administrator in the process. The next chapter outlines the design of the study and its participants.

CHAPTER THREE

Design and Methodology

Introduction

As a result of No Child Left Behind and the increased accountability on local school districts, it is important to understand how teacher evaluation can lead to increased student achievement. This study was designed to examine the relationship between teacher evaluation and student achievement.

Several factors presented a rationale for this study. The first was to determine the effectiveness of performance-based teacher evaluation. Districts must determine to what extent the teacher evaluation process is affecting student achievement. Much district time and money are spent on training teachers and administrators in current, research-based practices. There must be some way to determine the effectiveness of this investment in terms of student achievement. By examining the current practices of the teacher evaluation system of high-performing school districts, other school districts may

be able to glean practices that will benefit their children and staff.

Secondly, high-stakes testing leads to greater accountability with the public. Schools must show growth in student achievement on the state tests. With the increasing targets for No Child Left Behind, student achievement gains are paramount in keeping districts intact and providing viable options for families and communities.

And lastly, an examination of the law related to teacher evaluation will provide insight into current practices. School districts can make informed decisions to determine the best method of tying student achievement to teacher evaluation. The legal premise of teacher evaluation lends itself to interpretation.

Research Question

There was one question addressed in this study to conclusively answer the hypothesis.

1. What is the impact, if any, of using student achievement data to evaluate teacher performance on a performance-based teacher evaluation?

Subjects

The subjects used in this study were Missouri public school districts that were named on the Department of

Elementary and Secondary Education's "Top Ten Lists Based on MAP Test Results". The lists were retrieved from the Department of Elementary and Secondary Education's website and are normally accessible to the public. This list identified school buildings and districts whose performance ranked in the top ten based upon their percentages of students in the proficient and advanced levels on the Missouri Assessment Program. The subject areas of mathematics and communication arts were used. The lists used were from the 2005-06 and 2006-07 school year.

The school buildings and districts were ranked according to the number of times that they were listed in the "Top Ten" for those years. The buildings were also ranked by the number of students. The buildings were categorized as having 1) less than 250 students, 2) between 250 and 500 students, and 3) over 500 students. There were 133 school buildings identified in the "Top Ten" in the first category of less than 250 students. In the next category of 250 to 500 students, there were 131 school buildings identified on the list. For buildings with over 500 students, there were 88 buildings identified. For all three categories, there were 354

school buildings identified on the list. These buildings represented 181 different school districts.

Upon examination of the categories and the school buildings, there were many that were identified only one year for one subject area. The buildings identified more than one year and with more than one subject area or grade level were given preference in the study.

Sampling Procedure

Once the lists were compiled of the school districts and school buildings and the "Top Ten" were ranked in order of number of times each building was identified on the list, the top ten percent of the buildings were then identified for the study. A total of 45 school buildings were identified for the study. Buildings in different school districts were given preference in the study to allow for a larger number of districts to be included in the study.

Research Setting

The research setting included all schools that ranked on the Top Ten List for the MAP. The test administration of the Missouri Assessment Program subject areas is done with strict guidelines for all districts in the state of Missouri to follow. This lends to a standardized testing environment for the students. These

tests are administered by certified teachers or trained paraprofessionals within the given time constraints of the test administration. This setting is comparable in all districts in Missouri.

These assessments are scored by CTB-McGraw Hill and by trained scorers in Missouri. The selected response items are scored by CTB-McGraw Hill. Scoring sites are set up each summer to score the constructed response and performance event items from the Missouri Assessment Program assessments. The scorers must pass a rigorous training and scoring practice before they are allowed to score these assessments. All scores from CTB-McGraw Hill and the Missouri scorers are then sent to the Missouri Department of Elementary and Secondary Education where they are compiled for each student.

These results are then provided to each school district and parents of tested students in the fall annually. The school building and district- level scores are then ranked and published annually on the Missouri DESE website.

Research Design Procedure

This study examined the "Top Ten" schools ranked by the Department of Elementary and Secondary Education. These identified schools were surveyed, using a locally-

developed survey. The questions on the survey were asked by the researcher in an interview by telephone in the same manner and in the same order to each respondent to provide standardized conditions for the collection of data.

The primary purpose of this study was to determine if there is a relationship between the criteria specific to improving student achievement as one of the indicators on the teacher evaluation instrument to status on the Top Ten ranking. This can involve a discussion of student achievement during the evaluation process.

The dependent variable in this study included student results on the MAP test as indicated by the ranking of the schools on DESE's Top Ten list. The independent variable is the use of criteria on the performance-based teacher evaluation instrument that is specific to improving student achievement.

Data Analysis

Using the information obtained from the public website from the Department of Elementary and Secondary Education, http://www.dese.mo.gov/divimprove/assess/TopTen/, the researcher combined the Top Ten lists by size of schools for the 2005-06 and 2006-07 school years. The schools were then ranked based upon the possible

number of times that the building could be on the Top Ten list for communication arts and mathematics. For example, a school building that contains students in grades K-5 will have the opportunity to be on the Top Ten list for each of the two years in communication arts and mathematics for the third, fourth, and fifth grade levels. This means there are twelve opportunities for that school building to be listed on the Top Ten for those two years. A percent was calculated to determine how often the school building was identified in the Top Ten.

School buildings were also categorized by size. On the Top Ten lists, the Department of Elementary and Secondary Education group school buildings by number of students who attend the facility. The groupings include less than 250 students, between 250 and 500 students, and over 500 students.

A survey was conducted to determine whether or not a student achievement criterion was used in the teacher evaluation instrument. The principals responded to a series of five questions related to use of student achievement, observation, and any specific criterion related to student achievement. The results were tallied based upon the principals' responses and grouped

according to whether the response indicated a use of student achievement data in the teacher evaluation process.

Statistical Treatment of Data

Spearman's rank-order correlation coefficient is a measure of the linear relationship between two variables. It differs from Pearson's correlation only in that the computations are done after the numbers are converted to ranks. The data was converted to a table with the ranked-data values entered in a Spearman's test. This test is performed on data when there is a one within-subjects independent variable with two or more levels and a dependent variable that is not interval and normally distributed (but at least ordinal). This test is used to determine if there is a difference in the ranking of the scores. The null hypothesis in this test is that the distribution of the ranks of each type of score (i.e., communication arts and math on MAP) are the same.

The p value was measured to determine the reliability of the Spearman. This value indicated if the degree of statistical significance is valid. The results of the relationship between variables in the sample were considered less reliable the higher the p value.

Ethical and Political Considerations of the Study

While all DESE information was available to the general public through it website, the data gained from the telephone survey will remain confidential to the researcher and to the individual respondent. As a result, no personally-identifiable student information appears in this study. The results were shared with the respondents of the telephone survey through a summary sent via email.

Summary

Research concludes student achievement should be a topic of discussion during the performance-based teacher evaluation process. Data needed for this study was obtained from the Missouri Department of Elementary and Secondary Education website. The schools were ranked based upon the number of times they were in the Top Ten lists. The high-performing schools in Missouri were surveyed to determine if student achievement was a factor in teacher evaluations.

Great care was taken to ensure the confidentiality and reliability of the responses from the building administrators. District and school building names were removed from the data. In the following chapter, the

researcher will present the data gathered and analyze the results.

CHAPTER FOUR

Results

Introduction

This study investigated the relationship between current teacher evaluation practices and ranking of school buildings in the Top Ten list for Missouri schools. The researcher analyzed the ranking of those school buildings that used student achievement data as part of the teacher evaluation system with those that did not. This chapter presents the data relevant to teacher evaluation practices and increased student achievement.

Several factors were considered before the results of this study were examined. The data was taken from ranking the school buildings that were considered high-performing on the Missouri Assessment Program for the 2005-06 and 2006-07 school years. The school buildings were ranked according to the percentage of times that they could be ranked in the Top Ten for the subjects of communication arts and mathematics. The highest ranking buildings were considered in the study.

Results

Two years' worth of data was collected with school buildings ranked according to the percentage of times listed in the Top Ten ranking on the Missouri Assessment Program for communication arts and mathematics. The principals from the Top Ten schools responded to the question of using student achievement data during the teacher evaluation process. The ranking of the schools that responded positively were compared to the ranking of those that did not use student achievement data in the teacher evaluation process. The Spearman Rank order correlation coefficient was used to determine the significance of the relationship.

Table 1 depicts the rankings of the Top Ten with X representing the percent rank order of those schools that do use student achievement data in the teacher evaluation process and Y representing the rank order of those schools that do not use student achievement data in the teacher evaluation process. Table 2 depicts the rankings of the Top Ten schools. The X column represents those that use student achievement data and Y represents those that do not use student achievement data. There were ten schools in each category. The one- and two-tailed p value

was conducted to determine the reliability of the analysis of data.

Table 1

Spearman Rank Order Correlation Coefficient – Raw Data (Percent Top Ten) Ranking Used

	Percentage	
pairs	X	Y
1	83	75
2	81	75
3	75	60
4	75	42
5	70	42
6	63	42
7	58	33
8	42	33
9	42	25
19	33	25

n	rs	t	df
10	0.9596	9.65	8

	p value
1-tailed	0.000006
2-tailed	0.000011

Table 2

Spearman Rank Order Correlation Coefficient – Ranking Data Used

	Ranking	
pairs	X	Y
1	1	4
2	2	6
3	3	9
4	5	13
5	7	14
6	8	15
7	11	25
8	16	26
9	18	36
19	27	37

n	rs	t	df
10	0.9591	9.59	8

	p value
1-tailed	0.000006
2-tailed	0.000012

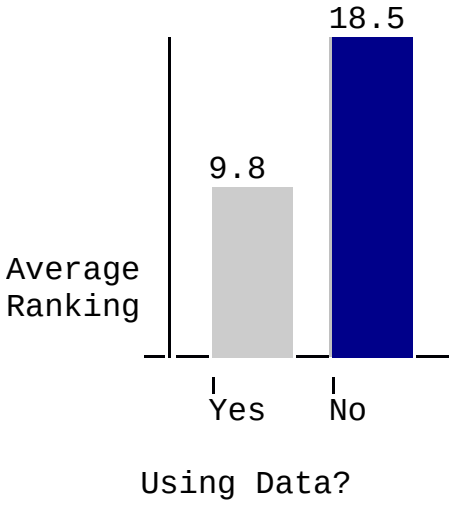

Figure 1 – Average Ranking in Top Ten for Schools Using Student Data vs. School Not Using Student Data in Teacher Evaluation Process

Analysis of Data

1. *Research Question Number One: What is the impact, if any, of using student achievement data to evaluate teacher performance during the performance-based teacher evaluation process?*

The Spearman was used to determine if a relationship exists between using student achievement data and ranking on the Top Ten in student performance. Table 1 results indicate there is a strong relationship with the Spearman correlation at .9596. This indicates a very strong relationship between the percent of times a school building is ranked in the Top Ten of student achievement and the use of student achievement data during the performance-based teacher evaluation process. This result is very valid with the p value being .000006.

The data was then converted to rankings and, again, the Spearman was used to determine if a relationship exists between the two factors – student achievement and ranking on the Top Ten of student achievement. Table 2 indicates another strong relationship using the Spearman with a .9591. The p value also indicates this is a strong relationship with .00006 on the one-tailed test.

Figure 1 depicts the average ranking on the Top Ten for student performance for each of the two categories,

those buildings that use student achievement and those that do not. There is an inverse relationship with the lower the number, the higher the ranking. The schools that use student achievement data during the teacher evaluation process had an average ranking of 9.8, while those that did not were ranked at 18.5. This, again, would indicate a strong relationship between using student achievement data during the performance-based teacher evaluation process and ranking on the Top Ten for student performance on the state assessment.

Deductive Conclusions

The results of this study revealed a significant correlation between using student achievement data on the performance-based teacher evaluation instrument and the ranking of those buildings on the Top Ten list for highest student performance on the state assessments. The level of the correlation is such that a strong relationship exists between the two factors. The literature review also provided studies over the past twenty years that would indicate the need to use student achievement data as part of the criteria to determine the effectiveness of teachers. These two, in conjunction with one another, lead a very strong argument for using student achievement data in teacher evaluation.

The null hypothesis:

1. There is no effect on student achievement when the performance-based teacher evaluation process specifically identifies and uses criteria related to improving student achievement.

The null hypothesis must be rejected based upon the results of the Spearman's Rank Order Correlation Coefficient. The results indicate a very strong relationship between using student achievement data and ranking on the Top Ten on the state assessment.

The statue in Missouri allows school districts to use the data by providing little guidance for school districts on the specific criteria that is to be used for teacher evaluation. These are left to the discretion of the school districts and Boards of Education.

Summary

The findings in this study were presented in this chapter. Chapter 5 will outline implications for school districts, as well as topics for further review and study. Recommendations for performance-based teacher evaluation processes will be discussed in the next chapter.

CHAPTER FIVE

Discussion

Introduction

Should teacher evaluation be tied to student achievement? How do public schools determine how effective a teacher is in the classroom? Those are two questions that all building principals must answer on a daily basis, especially in today's high-stakes world of education. This study was conducted to determine if a relationship exists between using student achievement data during the performance-based teacher evaluation process and the ranking on the Top Ten list for high student achievement. The specific criteria used on the teacher evaluation instrument can have an effect on student achievement. The findings from this study may allow administrators to re-evaluate the performance-based teacher evaluation criteria to determine which of those has the greatest impact on student achievement.

The following research question was examined to determine a relationship between using student

achievement data and ranking on the Top Ten in student performance on the state assessments.

1. What is the impact, if any, of using student achievement data to evaluate teacher performance during the performance-based teacher evaluation?

The rankings on the Top Ten in student performance on the state assessments were used to determine if a relationship exists between the process for teacher evaluation and increasing student achievement. This was done using the Spearman Rank Order Correlation Coefficient. The p-value was used to determine the strength of the validity of the relationship.

Implication for Effective Schools

The results of this study indicate a need to include student achievement data in the performance-based teacher evaluation process. This strong correlation indicated that using this data can strengthen the educational program for all students. The literature study further outlined the need and direction for this to occur.

School districts must be willing to do what is in the best interests of their students. Schools exist to educate children, not to provide employment for adults. By taking the results of this study and putting them into practice, schools can provide a quality education for

their students, as well as provide guidance for teachers in making good decisions for students.

Recommendations

The conclusions of this study provide several recommendations for further study. A larger population can be used to determine the strength of the study. This study was limited to the Top Ten schools that currently use student achievement data in the teacher evaluation process and match paired to those who do not that are in the Top Ten.

Schools can use the findings of this study as a model for improving student achievement in areas of concern. By focusing on student achievement data, school districts can provide an intensive measure to increase the areas of deficit for students. Teachers can then concentrate their efforts on the areas necessary for student success.

This study did not consider the wealth of the school district, in terms of assessed valuation or resources available to the educational staff. Further study can be done to break down the demographic data on the Top Ten schools to determine if wealth does make a difference in providing quality education for students.

The legal implications of using student achievement data are unwarranted and are used by different groups as reasoning not to use the student achievement data in the teacher evaluation process. Since this should not be a concern, districts can use the data to improve teacher effectiveness without legal ramifications.

Using best practices should provide a solid basis for improving student learning. The literature review provided case studies of student achievement success for others to emulate. Schools must find ways to increase student achievement that are research-based and provide good guidance for the students and staff.

Summary

School districts must be willing to step outside the historical restraints of teacher evaluation and determine what actually does make a difference for the students in their care. This study provides a comprehensive literature review on using student achievement data in the teacher evaluation process. Studies have been conducted and there are schools that use this data, successfully, when evaluating teachers. With the increasing accountability of public schools, the Boards of Education and educational community, as a whole, must

be willing to do whatever it takes to ensure the success of all students.

References

Attinello, J.R., Lare, D.W. & Source F. (2006). The value of teacher portfolios for evaluation and growth. NASSP Bulletin, 90(2), 132-152.

Bloom, B.S. (1981) *All our children learning,* McGraw-Hill, New York, NY.

Boyd, R. (1989). *Improving teacher evaluation*. Practical Assessment, Research & Evaluation, 1 (7). Retrieved February 12, 2009 from http://PAREonline.net/getvn.asp?v=1&n+7

Brandt, C. Mathers C., Oliva M., Brown-Sims M. & Hess J. (2007). Examining district guidance to schools on teacher evaluation policies in the Midwest Region (Issues & Answers Report, REL 2007-No. 030). Washington, DC: U. S. Department of Education. Retrieved January 28, 2009 from http://ies.ed.gov/ncee/edlabs/regions/midwest/pdf/REL_2007030_sum.pdf.

Braun, H.I. (2005) *Using student progress to monitor teachers: A primer on value-added models*. Princeton, NJ: Educational Testing Service. Retrieved January 28, 2009 from http://www.ets.org/Media/Research/pdf/PICVAM.pdf.

Brophy, J. & Good, T., (1970). Teacher's communication of differential expectations for children's classroom performance: Some behavior data. *Journal of Educational Psychology*, 61, 365-374.

CTB McGraw-Hill. (2007). Missouri assessment technical report. Retrieved November 28, 2007 from http://dese.mo.gov/divimprove/assess/index.html.

Danielson, C. (1996). Enhancing professional practice: A framework for teaching. Alexandria, VA: Association for Supervision and Curriculum Development and Educational Testing Service.

Danielson, C. & McGreal, T.L. (2000). *Teacher evaluation to enhance professional learning.* Princeton, NJ: Educational Testing Service.

Darling-Hammond, Linda. (2000). *Teacher quality and student achievement* [A Review of State Policy Evidence]. *Education Policy Analysis Archives, 8*(1), 1068-2341.

Denner, P.R., Miller, T.L., Newsome, J.D., & Birdsong, J.R. (2002). Generalizability and validity in the use of a case analysis assessment to make visible the quality of teaching candidates. *Journal of Personnel Evaluation in Education*, 16(3), 153-174.

DuFour, R., Eaker, R. & DuFour, R. (2005). *On common ground: The power of professional learning communities.* Bloomington, IN: National Educational Service.

Edmonds, R.R. (1982). Programs of school improvement: An overview. *Educational Leadership*, 40(3), 4-10.

Ellett, C. D. & Garland J. (1987). Teacher evaluation practices in our largest school districts: Are they measuring up to "state-of-the-art" systems? Journal of Personnel Evaluation in Education, 1(1), p. 69-92.

Fisher, D. & Frey, N. (2008). *Better learning through structured teaching: a framework for the gradual release of responsibility.*

Haefele, D.L. (1993) Evaluating teachers: a call for change. Journal of Personnel Evaluation in Education, 7(1), 21-31.

Holland, P., & Adams, P. (2002). Through the horns of a dilemma between instructional supervision and the summative evaluation of teaching. *International Journal of Leadership in Education,* 5(3), 227-247.

Marx, G. (2007). The principals' partnership research brief: teacher evaluation. Retrieved from www.principalspartnerhip.com/teacherevaluation.pdf.

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

Mathers, C., Oliva, M., with Laine, S. W. M. (2008). *Improving instruction through effective teacher evaluation: Options for states and districts.* Research and Policy Brief. Washington, DC: National Comprehensive Center for Teacher Quality.

Missouri Department of Elementary and Secondary Education. (1999) *Guidelines for performance-based teacher evaluation.*

Missouri Department of Elementary and Secondary Education. (2007a) *Missouri Assessment Program Technical Report 2007.* Retrieved February 18, 2008 from http://dese.mo.gov/divimprove/assess/tech/Final%20MAP%20Tech%20Report.pdf.

Missouri Department of Elementary and Secondary Education. (2007b). Missouri revised statutes. Retrieved November 30, 2007 from www.moga.mo.gov/statutes.

Missouri Department of Elementary and Secondary Education. (2005). Missouri school improvement

program. Retrieved December 5, 2007 from http://dese.mo.gov/divimprove/sia/msip%20overview.html.

Missouri Department of Elementary and Secondary Education. (2006). Missouri school improvement program standards and indicators manual: Accreditation standards for public school districts in Missouri. Retrieved November 5, 2007 from http://www.dese.mo.gov/divimprove/sia/msip/Fourth%20Cycle%20Standards%20and%20Indicators.pdf.

Missouri Department of Elementary and Secondary Education. (2008). Missouri school improvement program: understanding your annual performance report 2008-09. Retrieved November 12, 2008 from http://www.dese.mo.gov/divimprove/sia/dar/understandingyourAPR.pdf

Pollock, Jane E. (2007). *Improving student learning one teacher at a time,* Alexandria, VA: Association of Supervision and Curriculum Development.

Ravitch, Diane (2007) *EdSpeak*, Alexandria, VA: Association of Supervision and Curriculum Development.

Sawa, R., (1995). *Teacher evaluation policies and practices.* Retrieved November 20, 2008 from

http://saskschoolboards.ca/research/instruction/95-04.htm#toc.

Schmocker, Michael (2006) *Results now: how we can achieve unprecedented improvements in teaching and learning,* Alexandria, VA: Association of Supervision and Curriculum Development.

Shannon, D.M. (1991, February). *Teacher evaluation: A functional approach*. Paper presented at the annual meeting of the Eastern Education Research Association, Boston.

Stanley, S.J. and Popham, J.W. (1988). Teacher evaluation: six prescriptions for success. Alexandria, VA: Association for Supervision and Curriculum Development.

Stiggins, R.J. (1995). Professional development: The key to a total quality assessment environment. *NASSP Bulletin*, 79 (573), 11-19.

Stronge, J.H. (2007). Qualities of effective teachers (2nd ed). Alexandria, VA: Association for Supervision and Curriculum Development.

Toch, T. & Rothman, R. (2008). R*ush to judgment: Teacher evaluation and teacher quality.* Retrieved February 12, 2009 from www/readingrockets.org/article/29033?theme=print.

Tucker, Pamela D. & Stronge, James H. (2005) *Linking teacher evaluation and student learning*, Alexandria, VA: Association of Supervision and Curriculum Development.

U.S. Department of Education. (2000). *NAEP 1999 Trends in Academic Progress: Three Decades of Student Performance,* NCES 2000-469, by J.R. Campbell, C.M. Hombo, and J. Mazzeo. Office of Educational Research and Improvement. National Center for Education Statistics. Washington, DC: 2000.

United States Department of Education. (2004). No child left behind: A toolkit for teachers. Retrieved October 20, 2007 from www.ed.gov/teachers/nclbguide nclb-teachers-toolkit.pdf.

United States Department of Education (2007). *The nation's report card 2007 at a glance*. Retrieved February 20, 2009 from http://nces.ed.gov/nations Reportcard/pdf/about/2009486.pdf.

Vescio, V., Ross, D. & Adams, A. (January 2006). *A review of research on professional learning communities: What do we know?* NSRF Research Forum: University of Florida.

Weiss, E.M. & Weiss, G. (1998). *New directions in teacher evaluation*. Washington D.C.: ERIC Clearinghouse of

Teaching and Teacher Education. (ERIC Document Reproduction Service No. ED429052)

Whitehurst, G. (2002). Research on teacher preparation and professional development. Washington D.C.: White House Conference on Preparing Tomorrow's Teachers, 2001.

Wittrock, M. C. (editor). (1986). *Handbook of research on teaching.* 3rd edition. New York: Macmillan Publishing Co.

Appendix A

Missouri Statute Section 168.128:

Teacher records, how maintained-evaluations, how performed and maintained.

-The board of education of each school district shall maintain records showing periods of service, dates of appointment, and other necessary information for the enforcement of section 168.102 to 168.130. In addition, the board of education of each school district shall cause a comprehensive performance-based evaluation for each teacher employed by the district. Such evaluation shall be ongoing and of sufficient specificity and frequency to provide for demonstrated standards of competency and academic ability. All evaluations shall be maintained in the teacher's personnel file at the office of the board of education. A copy of each evaluation shall be provided to the teacher and appropriate administrator. The State Department of Elementary and Secondary Education shall provide suggested procedures for such an evaluation.

(L. 1969 p.275§168.114, A.L. 1983 H.B. 38 & 783)

Appendix B

Missouri Performance-Based Teacher Evaluation Criteria:

Standard 1: The teacher causes students to actively participate and be successful in the learning process.

Criteria for Standard 1: The teacher…

1. causes students to acquire the knowledge and skills to gather, analyze and apply information and ideas.
2. causes students to acquire the knowledge and skills to communication effectively within and beyond the classroom.
3. causes students to acquire the knowledge and skills to recognize and solve problems.
4. causes the students to acquire the knowledge and skills to make decisions and act as responsible members of society.

Standard 2: The teacher uses various forms of assessment to monitor and manage student learning.

Criteria for Standard 2: The teacher…

1. causes various ongoing assessment to monitor the effectiveness of instruction.
2. provides continuous feedback to students and family.
3. assists students in the development of self-assessment skills.
4. aligns the assessments with the goals, objectives, and instructional strategies of the district curriculum guides.
5. uses assessment techniques that are appropriate to the varied characteristics and developmental needs of students.

Standard 3: The teacher is prepared and knowledgeable of the content and effectively maintains students' on-task behavior.

Criteria for Standard 3: The teacher…

1. demonstrates appropriate preparation for instruction.
2. chooses and implements appropriate methodology and varied instructional strategies that address the diversity of learners.

3. creates a positive learning environment.
4. effectively manages student behaviors.

Standard 4: The teacher communicates and interacts in a professional manner with the school community.

Criteria for Standard 4: The teacher…

1. communicates appropriately with students, parents, community and staff.
2. engages in appropriate interpersonal relationships with students, parents, community, and staff.

Standard 5: The teacher keeps current on instructional knowledge and seeks and explores changes in teaching behaviors that will improve student performance.

Criteria for Standard 5: The teacher…

1. engages in professional development activities consistent with the goals and objectives of the building, district and state.
2. engages in professional growth.

Standard 6: The teacher acts as a responsible professional in addressing the overall mission of the school district.

Criteria for Standard 6: The teacher…

1. adheres to all the policies, procedures, and regulations of the building and district.
2. assists in maintaining a safe and orderly environment.
3. collaborated in the development and/or implementation of the district's vision, mission, and goals.

Table A1

Timeline of the Study

Date	Event
Spring 2006	Missouri MAP tests administered to all 3-8, 10, and 11 grade students
March 2007	Top Ten List published for 2005-06 scores
Spring 2007	Missouri MAP tests administered to all 3-8, 10, and 11 grade students
March 2008	Top Ten List published for 2006-07 scores
Fall 2008	Top Ten Lists compiled
Winter 2008	Data gathered and analyzed for the purpose of the study.

Table A2

Ranking of Buildings by Percentage for Top Ten

	School	Enrollment	2006	2007	Total	Percentage	Yes/No
1	A1	774	5	5	10	83%	Yes
2	A2	641	3	5	8	81%	Yes
3	A3	555	6	3	9	75%	Yes
4	A4	598	5	3	8	75%	No
5	B1	297	6	6	12	75%	Yes
6	C1	108	6	3	9	75%	No
7	A5	1096	6	8	14	70%	Yes
8	B2	338	6	4	10	63%	Yes
9	A6	1548	6	6	12	60%	No
10	A7	532	4	3	7	58%	No
11	A8	537	5	3	8	50%	Yes
12	C2	137	5	6	11	46%	No
13	A9	569	1	4	5	42%	No
14	B3	282	3	2	5	42%	No
15	B4	299	2	3	5	42%	No
16	B5	284	2	3	5	42%	Yes
17	B6	351	1	4	5	42%	No
18	C3	108	3	2	5	42%	Yes
19	C4	120	2	3	5	42%	No
20	B7	256	3	3	6	38%	No
21	B8	334	2	1	3	38%	No
22	C5	118	2	4	6	38%	No
23	C6	106	2	4	6	38%	No
24	C7	75	3	3	6	38%	No
25	A10	525	4		4	33%	No
26	A11	677	1	2	3	33%	No
27	A12	723	1	2	3	33%	Yes
28	B9	373	2	6	8	33%	No
29	B10	258	3	1	4	33%	No
30	B11	265	2	2	4	33%	No
31	B12	256	2	2	4	33%	No
32	B13	453		2	2	33%	No
33	C8	188	2	2	4	33%	No
34	C9	131	1	3	4	33%	No
35	C10	166	4	1	5	31%	No
36	A13	534	2	1	3	25%	No

37	A14	559	1	1	2	25%	No
38	B14	292	1	1	2	25%	No
39	C11	116	1	4	5	25%	No
40	C12	75	1	3	4	25%	No
41	C13	189	2	2	4	25%	No
42	C14	129	2	2	4	25%	No
43	C15	165	1	3	4	25%	No
44	C16	137	2	2	4	25%	No
45	C17	129	2	3	5	21%	No

Table A3

Top Ten Buildings Ranked by & Possibility in Top Ten

School Buildings with Enrollment over 500

School	Enrollment	2006	2007	Total	Percentage
A1	774	5	5	10	83%
A2	641	3	5	8	81%
A3	555	6	3	9	75%
A4	598	5	3	8	75%
A5	1096	6	8	14	70%
A6	1548	6	6	12	60%
A7	532	4	3	7	58%
A8	537	5	3	8	50%
A9	569	1	4	5	42%
A10	525	4		4	33%
A11	677	1	2	3	33%
A12	723	1	2	3	33%
A13	534	2	1	3	25%
A14	559	1	1	2	25%

School Buildings with Enrollment between 250 and 500

School	Enrollment	2006	2007	Total	Percentage
B1	297	6	6	12	75%
B2	338	6	4	10	63%
B3	282	3	2	5	42%
B4	299	2	3	5	42%
B5	284	2	3	5	42%
B6	351	1	4	5	42%
B7	256	3	3	6	38%
B8	334	2	1	3	38%
B9	373	2	6	8	33%
B10	258	3	1	4	33%
B11	265	2	2	4	33%
B12	256	2	2	4	33%
B13	453		2	2	33%
B14	292	1	1	2	25%

School Buildings with Enrollment Less than 250

School	Enrollment	2006	2007	Total	Percentage
C1	108	6	3	9	75%
C2	137	5	6	11	46%
C3	108	3	2	5	42%
C4	120	2	3	5	42%
C5	118	2	4	6	38%
C6	106	2	4	6	38%
C7	75	3	3	6	38%
C8	188	2	2	4	33%
C9	131	1	3	4	33%
C10	166	4	1	5	31%
C11	116	1	4	5	25%
C12	75	1	3	4	25%
C13	189	2	2	4	25%
C14	129	2	2	4	25%
C15	165	1	3	4	25%
C16	137	2	2	4	25%
C17	129	2	3	5	21%

Appendix C

Survey

School Building/District:
Contact Person:

Hello, this is Lucy Lyon. I am currently a doctoral student with Lindenwood University in St. Charles. I am gathering my statistical data for my dissertation. My topic is Teacher Evaluation and Student Achievement. Your school was chosen to participate due to its status ranking in the Top Ten Performing schools on the MAP tests for 2006 and 2007. I would appreciate a few minutes of your time to gather information regarding your teacher evaluation practices and its relationship to the high performance of your students on the MAP.

1. On your Teacher Evaluation Instrument, is student achievement a formal part of the process? Is there a specific criterion tied to student achievement on standardized tests/common assessments?

2. If so, how data does a teacher show that this criterion has been met?

3. What is the exact wording on your instrument for the criteria related to increasing student achievement?

4. Are there a specific number of formal observations required annually? (A formal observation is one that creates a document after the observation) If so, how many?

5. Are there a specific number of informal observations required annually for each teacher? (Walk-through with no documentation)? If so, how many?

6. Would it be possible to get copy of your Teacher Evaluation Instrument? Email or fax or website?

7. Would you like a copy of the results of this survey? Email address?

Appendix D

LINDENWOOD UNIVERSITY

Application for IRB Review of
Research Proposal Involving Human Subjects

1. Title of Project: Project # _____
(To be filled out by IRB chairman)

TEACHER EVALUATION PRACTICES AND STUDENT ACHIEVEMENT

2. Faculty Advisor: Department:
Dr. Terry Reid Education

Extension: Email:
417-881-0009 treid@lindenwood.edu

3. Primary Investigator(s): Department:
Lucy Lyon

Local Phone: Email:
417-678-4918 llyon@hdnet.k12.mo.us

4. Anticipated starting date for this project:
Fall 2008

5. Anticipated ending date for this project:
April 2009

6. State the hypothesis of the proposed project:

1. Is there an effect on student achievement when the teacher evaluation system specifically identifies criteria related to improving student achievement? Is there discussion of student achievement during the teacher evaluation process?

7. State the purpose (objectives) and rationale of the proposed project. Include any questions to be investigated.

The purpose of this study is to determine the effect of teacher evaluation systems on student achievement. When school personnel are held accountable

for student achievement through an evaluation system, does student achievement increase? Teacher accountability can focus the efforts on actual teaching performance and provide a structure to improve the teaching abilities of staff. Effective teaching must be defined, measured, and related to student achievement. The following research questions will be addressed:

1. What is the impact, if any, of using student achievement data to evaluate teacher performance during the performance-based teacher evaluation process?

8. Has the research project been reviewed or is it currently being reviewed by an IRB at another institution? If so, please state when, where and disposition (approval/non-approval/pending).

The research project has not or is not currently being reviewed by an IRB at another institution.

9. Participants involved in the study:

a. Indicate how many persons will be recruited as potential participants in this study.

LU participants	__0__	Undergraduate
	__0__	Graduate
	__0__	Faculty and/or staff
		Non LU participants
	__0__	Children
	__0__	Adolescents
	__20_	Adults
	__0__	Seniors
	__0__	Persons in institutional settings (e.g. nursing homes, correctional facilities)

Other (specify):

b. From what source will the potential participants be recruited? N/A

_____ LU undergraduate and/or graduate classes
_____ LU Human Subject Pool (LUHSP)
__X_ School boards (districts)
_____ Greater St. Charles community
_____ Agencies (please list)_____
_____ Businesses (please list)_____

____ Health care settings, nursing homes, etc.(please list)_____

Other (specify):

Department of Elementary and Secondary Education website

c. If any persons within the selected group(s) are being excluded, please explain who is being excluded and why. (Note: According to the Office of LUHSP, all students within the LU Human Subject Pool must be allowed to participate, although exclusion of certain subjects may be made when analyzing data.)

School districts that are not in the Top Ten performing schools of the state based upon their Missouri Assessment Program scores will not be considered for this study.

d. Describe how and by whom the potential participants will be recruitment (e.g. poster, flyers, advertisements, letters, telephone and other verbal scripts).

School districts that score in the Top Ten performing schools of the state based upon the Missouri Assessment Program will be recruited for the study by email and telephone.

e. Where will the study take place?

____ On Campus – Explain:

_X__ Off Campus- Explain:

Southwest Missouri, Lindenwood University

10. Methodology/procedures

a. Provide a sequential description of the procedures to be used in this study.

1. Determine the school districts that are in the Top Ten of scoring on the Missouri Assessment Program
2. Gather the Teacher Evaluation Document from each of these districts
3. Through either telephone interview or email questionnaire, determine if the Teacher Evaluation Instrument meets the selected criteria for the study.
4. Conduct a Spearman's Rank Order Correlation Coefficient

b. Which of the following procedures will be used? Provide a copy of all materials to be used in the study.
____ Surveys or questionnaires (mail back)- Are they standardized?
_X__ Surveys or questionnaires (in person)- Are they standardized? No
____ Computer administered task or survey- Are they standardized?
____ Interviews (in person)
_X__ Interviews (by telephone)
____ Focus groups
____ Audiotaping
____ Video Taping
_X__ Analysis of secondary data (no involvement with human participants)
____ Invasive physiological measurement (e.g. venipunture, catheter insertion, muscle biopsy, collection of other tissues, etc.) Explain:
____ Other (Specify)

11. How will the results of this research be made accessible to participants? Explain and attach a copy of any forms used.

Results of the study will be shared with participants upon request via email.

12. Potential Benefits and Compensation for the Study:
a. Identify and describe anticipated benefits (health, psychological or social benefits) to the participants from their involvement in the project.

Results of the study will be shared with the participants

b. Identify and describe any known or anticipated benefits to society from this study.
The results of this study can be duplicated within school districts to determine if use of specific criteria related to student achievement on the Teacher Evaluation instrument increases student achievement.

c. Describe any anticipated compensation (monetary, grades, extra credit, other) to participants.

There is no anticipated compensation to participants.

13. Potential Risks from the Study:

a. Identify and describe any known or anticipated risks to participants involved in this study. Include physiological, emotional, social, economic, legal, etc. risks/stressors. A study specific medical screening form must be included when physiological assessments are used and associated risks to participants are greater than what would be expected in normal daily activities.

There are no anticipated risks to participants involved in this study.

b. Will deception be used in study? If so explain rationale.

Deception will not be used in this study.

c. Does this project involve information about sensitive behavior, such as sexual behavior, drug/alcohol abuse, or illegal behavior? If so explain

This study does not involve information about sensitive behavior.

14. Informed Consent Process:

a. What process will be used to inform the potential participants about the study details and to obtain their consent for participation?

_X_ Information letter with written consent form for participants or their legally authorized agents; provide a copy (via email).
___ Information letter with written or verbal consent from director of institutions involved; provide a copy.
____ Information letter with written or verbal consent from teachers in classrooms or daycare; provide a copy
Other (specify):

b. What special provisions have been made for informed consent for non-English speaking persons, mentally disabled or other populations for whom there may be difficulty in providing informed consent?

If necessary, special provisions (interpreters, native language documents, etc.) will be made for informed consent for non-English speaking persons, mentally disabled or other populations for whom there may be difficulty in providing informed consent.

15. Anonymity of Participants and Confidentiality of Data:

a. Explain the procedures to be used to ensure anonymity of participants and confidentiality of data both during the research and in the release of the findings.

   No names or identifying information will be used in the analysis or results of the study. Participants will be assigned random numbers for inclusion in the study.

b. How will confidentiality be explained to the participants?

   Confidentiality will be explained to the participants through the letter of consent and email.

c. Indicate the duration and location of secure data storage and the method to be used for final disposition of the data.

   Data will be stored in a locked filing cabinet and shredded after three years.

Paper Records

__X___ Confidential shredding after _3___years
_____ Data will be retained indefinitely in a secure location.
_____ Data will be retained upon completion of specific course and then destroyed.

Audio/video Recordings

_____ Erasing of audio/video tapes after ____years

_____ Data will be retained indefinitely in a secure location.

_____ Data will be retained upon completion of specific course and then destroyed.

Electronic Data

_____ erasing of audio/video tapes after ____years

_____ Data will be retained indefinitely in a secure location.

_____ Data will be retained upon completion of specific course and then destroyed.

Other:

Specify Location:

16. Researchers must ensure that all supporting materials/documentation for their applications are submitted with the signed, hard copies of the IRB Research Proposal Form. Please check below all appendices that are attached as part of your application package. Submission of an incomplete application package will increase the duration of the IRB review process.

_____ Recruitment materials: A copy of any posters, fliers, advertisement, letters, telephone or other verbal scripts used to recruit/gain access to participants (see 9d).

_____ Materials: A copy of all surveys, questionnaires, interview questions, interview themes/sample questions for open-ended interviews, focus group questions, or any standardized tests used to collect data (see 10b).

__X__ Feedback letter (see 11). – Via Email

_____ Medical screening form: Must be included for all psychological measurements involving greater than minimal risk, and tailored for each study (see 13 a).

_____ Information letter and consent forms used in studies involving interaction with participants (see 14a).

_____ Information/cover letter used in studies involving surveys or questionnaires (see 14a.)

_____ Parent information letters used in studies involving surveys or questionnaires (see 14a).

_____ Other:

VITA

Lucy K. Lyon was born April 14, 1961 in Aurora, Missouri. She graduated from Aurora High School in 1979. After this, she earned her bachelor of science degree in elementary education from Missouri State University (1982), a master of science in education degree from Missouri State University (1988), an educational specialist degree from Lindenwood University (2008), and doctorate degree in educational leadership from Lindenwood (2009).

Lucy served as a teacher for the Aurora R-VIII School District, Aurora, Missouri, for eleven years and as an elementary principal for the Crane R-III School District, Crane, Missouri, for ten years. She has then served as assistant superintendent for the Aurora R-VIII School District, Aurora, Missouri for the past five years.

Lucy has two daughters, Regina and Erin.