IMPLEMENTATION OF TEACHER EVALUATION MANDATES IN THE STATE OF MICHIGAN: HOW PRINCIPALS PERCEIVE THE NEW MANDATES AND ITS IMPACT ON LOCAL IMPLEMENTATION

Sarah A. Starr-Gwinn

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Education

Department of Educational Leadership

Central Michigan University
Mount Pleasant, Michigan
February 2014

ACKNOWLEDGEMENTS

I wish to thank members of my dissertation committee: Dr. Regina Umpstead, Dr. Frim Ampaw, and Dr. Derek Anderson. These faculty members provided direction, feedback, and support throughout the entirety of the dissertation process. Their valued expertise is truly appreciated.

As dissertation chair, Dr. Umpstead provided her esteemed knowledge on the topic of teacher evaluation law in Michigan and the continual changes involved with evaluation reform. Dr. Umpstead challenged my thinking to encourage a more in-depth look at the topic of study while providing encouragement throughout the dissertation process.

ABSTRACT

IMPLEMENTATION OF TEACHER EVALUATION MANDATES IN THE STATE OF
MICHIGAN: HOW PRINCIPALS PERCEIVE THE NEW MANDATES AND ITS IMPACT
ON LOCAL IMPLEMENTATION

by Sarah Starr-Gwinn

This non-experimental quantitative study was conducted to measure principals'
perceptions of the new teacher evaluation mandates in the state of Michigan. The study aimed to
assess principals' experiences and perceptions in response to the new teacher evaluation
mandates to determine if perceptions, position, type of school, years of service, prior teaching
experience, evaluation training, and gender impacted local teacher evaluation implementation in
public schools across Michigan.

At the local level, principals are largely responsible in leading the implementation
process. Principals are accountable for proper local implementation of the newly defined teacher
evaluation system. With evaluation results tied to high-stakes decisions, it is important to
develop a deeper understanding of local districts' experiences in the execution of the new
evaluation system.

This quantitative study surveyed elementary, middle, and high school principals and vice-
principals in public schools in Michigan. A survey was developed based on evaluation research,
the Michigan Council for Educator Effectiveness Interim Report, and through consultation with
the dissertation committee. The survey was developed to assess principals' perception of the
new mandates and the implementation of the new mandates within local districts. A pilot study
was conducted to increase survey clarity and organization and revisions were made based on the
pilot study recommendations prior to survey administration. The survey was administered to
1200 public school principals and vice principals at the elementary, middle, and high school

level in Michigan. Participants were chosen using a stratified random sample to ensure equal representation of elementary, middle, and high school principals and vice-principals. The survey was administered via e-mail through Qualtrics and survey data were transferred to SPSS for analysis.

This study sought to determine whether principals' prior experiences influenced their perceptions in regard to the newly defined teacher evaluation mandates and the implementation of the mandates within their local school district. The findings indicated that years of service, type of school, prior teaching experience, evaluation training, and gender did not have a significant impact on a principal's perceptions of the new evaluation mandates. Different groups of principals did not indicate differences in perceptions of the new mandates, meaning that perceptions of the mandates were generally consistent within the study. Principals were surveyed regarding implementation of both formal and informal observation methods in the evaluation process. Analysis of results indicated years of service, position, type of school, perceptions, prior teaching experience, evaluation training, and gender did not have a significant impact on the implementation of the new mandates.

Although the findings of this study are specific to the public schools within Michigan, teacher evaluation is a common measure of reform across school districts in the United States. Many states have implemented new systems for evaluation with increased accountability based on evaluation results. It is crucial that continued research efforts aim to develop more comprehensive evaluation systems with a focus on the factors that impact implementation in local districts.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Overview

The defined needs of society across the nation have impacted the educational system, the

defined purpose of education, and the culture of schools throughout history and continue to

influence the system today. The current climate in the United States has created a sense of

urgency once again in regard to educational reform and this has led to increased demands in

response to a belief that our nation's schools are failing. "Policy makers, parents, and the public

want tangible proof that schools are getting the job done" (Deal & Peterson, 2009, p. 69) and

national and state governments have passed legislation to increase accountability in schools

across the nation. With the increase in external accountability and state mandates, local school

districts are responsible for implementation of the new reform measures.

Due to the increased demands across the United States and current legislation within the

state of Michigan, districts are required to implement a more rigorous, transparent, and fair

teacher evaluation process to include observation tools and evaluations that are examined for

reliability and validity. As a result, school districts and principals are responsible for the

implementation of the new teacher evaluation mandates at the local district level. A change to

Michigan state law is responsible for the change in mandates, yet this does not guarantee

consistent implementation across the local school districts. This study assessed principals'

perceptions and experiences in response to the new teacher evaluation mandates to determine if a

principal's years of service, position, type of school, gender, prior teaching experience,

evaluation training, and perceptions impacted local implementation in public school districts

across Michigan.

Background

The various stakeholders in education provide for increased demands and changes in expectations for districts, schools, educators, and students. The most recent educational reform has resulted in legislation and competitive grant programs created by the federal government to encourage increased accountability in our nation's schools. The federal government has passed legislation to increase the nation's investment in education, with competitive grant programs such as Race to the Top, designed to encourage educational innovation and reform at the state level (United States Department of Education, 2009, p. 2). President Barack Obama also introduced The Blueprint for Reform (2010), as the reauthorization of the Elementary and Secondary Education Act of 1965 (ESEA), as "not only a plan to renovate a flawed law, but also an outline for a re-envisioned federal role in education" (United States Blueprint for Reform, 2010, p. 2). The Blueprint for Reform (2010) "builds on the significant reforms already made in response to the American Recovery and Reinvestment Act (ARRA) of 2009" (United States Department of Education, 2010, p. 3). The Blueprint for Reform incorporated the framework of the ARRA and extended the framework with a focus on the key educational priorities: college- and career-ready students, great teachers and leaders in every school, equity and opportunity for all students, raising the bar and rewarding excellence, and promoting innovation and continuous improvement.

President Obama signed into law the American Recovery and Reinvestment Act of 2009 (ARRA) designed to stimulate the economy, encourage job creation, with an increased focus on investments in critical sectors including education. The ARRA (2009) provided $4.35 billion for the Race to the Top Fund, "a competitive grant program designed to encourage and reward States

that were creating the conditions for education innovation and reform" (U.S. Department of Education, 2009, p. 1). The Race to the Top grant rewarded States that had:

- Created the conditions for education innovation and reform.

- Achieved significant improvement in student outcomes; including making substantial gains in student achievement, closing achievement gaps, improving high school graduation rates, and ensuring student preparation for success in college and careers.

- Implemented ambitious plans in four core education reform areas:

  o Adopting standards and assessments that prepare students to succeed in college and the workplace and to compete in the global economy.

  o Building data systems that measure student growth and success, and inform teachers and principals about how they can improve instruction.

  o Recruiting, developing, rewarding, and retaining effective teachers and principals, especially where they are needed most.

  o Turning around low achieving schools. (p. 2)

The Race to the Top competitive grant program was designed to encourage and reward states that demonstrated educational innovation and reform plans within their state with a focus on increasing student achievement. Several criteria were outlined with the grant program in order to choose states that demonstrated innovation and reform and had clear plans in place for future reform. Michigan applied for the Race to the Top competitive grant program, but was not chosen as a finalist.

One of the selection criteria focused on the use of data systems to support instruction and another area of criteria focused on providing great teachers and leaders in all schools with the

goal of improving instruction and student achievement. States were encouraged to develop and use longitudinal data systems to support and improve instruction and guide the development of effective professional development. Professional development and the use of data systems would also promote the effort in assuring that great teachers and leaders were in all schools. A main focus within this criteria included improving teacher and principal effectiveness based on performance through effective teacher evaluation. These criteria included several elements that are evident within the Michigan teacher evaluation reform process:

> Design and implement rigorous, transparent, and fair evaluation systems for teachers and principals that (a) differentiate effectiveness using multiple rating categories that take into account data on student growth as a significant factor and (b) are designed and developed with teacher and principal involvement. (Race to the Top, 2009, p. 2)

In an effort to promote conditions for innovation and reform, Race to the Top (2009) encouraged schools to adopt standards and assessments that prepare students for success in college and to compete in a global economy. In response to The Race to the Top monetary incentive, states created statewide mandates and expectations to increase accountability through the teacher evaluation processes. Although Michigan was not named a finalist in the second round of the competitive Race to the Top grant, the state planned to move forward with educational reform initiatives.

Reform in Michigan

Governor Jennifer Granholm signed a five-bill legislative package in January 2010, with the goal of reforming Michigan's educational system. One main concentration of the legislation focused on improving the teacher evaluation process, and connecting student performance as a component in teacher and administrator evaluation. According to the State of Michigan

Education Reform Legislative Analysis (2010), the legislation required school boards to adopt a rigorous, transparent, and fair performance evaluations. To do this, districts had to establish clear approaches for measuring student growth with growth as a significant factor in the evaluation of teachers' performance. Although, the State of Michigan mandated an improved teacher evaluation process across the state, local districts were responsible for the development of the performance evaluation. At the time the current law and policies were unclear and unhelpful to districts in the process of developing a more rigorous system. "Leaders pushed the problem on the budget-strapped local districts, many of which lack the capacity, expertise and resources to do it (The Education-Trust Midwest, 2011, p.2). As districts began this process, new expectations and mandates continued to evolve at the state level.

With the election of Governor Snyder and several new legislators in Michigan, accountability was again at the forefront with a continued focus on teacher evaluation reform. The new legislature passed Public Act 102 (2011) that included amendments to the previous evaluation law. Section 1249 was amended to include criteria that school districts must adopt and implement a performance evaluation system that evaluates the teacher's job performance at least annually, while providing timely and constructive feedback. Districts were required to implement these new evaluation systems no later than September 2011.

The passage of the new legislation during the time frame when districts were implementing the initial changes to their teacher evaluation systems complicated their development and implementation. For the 2011-2012 and 2012-2013 school year, districts could implement their locally determined educator evaluation system for all educators, using Michigan Department of Education data when appropriate. The performance evaluation should have included the following criteria: evaluate the teacher's job performance at least annually, use

5

multiple rating categories that take into account student growth with timely and constructive feedback, establish clear approaches to measuring student growth while providing teachers with relevant data on student growth, and districts must use the evaluation to inform decisions regarding promotion, retention, and development of teachers. Starting at the end of the 2012 school year districts were required to report effectiveness ratings for all teachers based on four ratings: highly effective, effective, minimally effective, and ineffective.

During the first two years of implementation, districts were allowed to develop and implement their own unique systems, as long as they met the basic requirements defined by the state. Following the first two years of implementation, 50% of Michigan school districts reported using Charlotte Danielson's Enhancing Professional Practice for Performance of Teaching as the framework for their evaluation systems (Keesler & Howe, 2012). The second highest framework identified by school districts was a combined measure including the Danielson framework and a local component (Keesler & Howe). During the first and second year of implementation of the law "Michigan had over 800 unique district evaluation systems," making any comparison of reported effectiveness ratings difficult due to an inconsistent standard rigor across districts (Keesler & Howe, 2012, p. 1).

In response to Michigan legislation, Michigan's governor and legislative members appointed a council to work to develop common tools and expectations for local school-district implementation starting in 2013. The Michigan Council for Educator Effectiveness (MCEE) was established with council members appointed in September 2012, to include five voting members (Executive order, 2012). Governor Snyder appointed Deborah Loewenberg Ball, dean of the University of Michigan School of Education, as chair of the MCEE. Snyder also appointed Mark Reckase from Michigan State University's College of Education, and Nick Sheltrown from

National Heritage Academy in Grand Rapids as council members. Senator Majority Leader

Randy Richardville appointed David Vensel, a principal from Jefferson High School in Monroe,

and Speaker of the House Jase Bolger appointed Jennifer Hammond, a principal from Grand

Blanc. Joseph Martineau, representative from the Michigan Department of Education, also

served on the MCEE as a non-voting member. Public Act 102 (2011) also established the

Advisory Committee to the MCEE, consisting of Governor-appointed teachers, members of

interest groups, and district leaders responsible for answering questions submitted by the council

and providing feedback on the observation and student growth components.

　The MCEE was created as a two-year temporary agency to develop a "fair, transparent,

and feasible" evaluation system for teachers and administrators based on rigorous standards of

professional practice and measurement (MCEE, 2012, p. 1). The MCEE (2013) completed its

work through two primary processes:

> (1) conducting a pilot study of educator effectiveness tools to provide data on
>
> implementation and validity and crucial feedback from education professionals; and (2)
>
> meetings, consultation, and research. (p. 4)

The MCEE conducted a year-long pilot study, Pilot of Educator Effectiveness Tools,

in thirteen districts across Michigan to gather data to inform the MCEE final recommendations.

The pilot study team provided the MCEE with updates, insight, and challenges presented

throughout the pilot study based on feedback from administrators and teachers involved. The

council met regularly, twice a month in whole group meetings, as well as established subgroups

which met to work on high priority items between council meetings. Meetings were also

regularly scheduled to allow for stakeholder and expert input throughout the process.

The suggested framework established by the MCEE called for more meaningful evaluations that are completed in a timely, transparent, and efficient manner. The MCEE (2012) defined the common vision to ground the efforts of the council:

> The Michigan Council for Educator Effectiveness will develop a fair, transparent, and feasible evaluation system for teachers and school administrators. The system will be based on rigorous standards of professional practice and of measurement. The goal of this system is to contribute to enhanced instruction, improve student achievement, and support ongoing professional learning. (p. 1)

The Michigan Council for Educator Effectiveness (2012) indicated that prior to full implementation of the newly designed teacher evaluation observation tool, a pilot-testing year would be imperative beginning in the fall of 2012. The intent of the pilot year was to allow the state to pilot a set of recommended tools in a small number of districts to learn how well each worked and to identify problems and technical issues prior to full implementation in the state. The council then reviewed all recommendations and made adjustments based on the new data and information following the pilot year. In the fall of 2013 the MCEE released the final recommendations for a statewide teacher evaluation system: building an improvement-focused system of educator evaluation in Michigan (MCEE, 2013). The MCEE also indicated a necessary phasing-in period to give evaluators the opportunity to become trained in using the observation tool to increase fidelity, validity, and reliability.

The MCEE is still in the process of gathering and analyzing data based on the pilot year findings in regard to the teacher evaluation frameworks piloted. Districts may still choose to implement their own evaluation system at the local level, but the observation tool must be rigorously examined for reliability and validity, as well as meet the criteria determined at the

8

state level. Districts are required to apply for a waiver from the state system in order to continue implementation of their own local evaluation tools (MCEE, 2013).

According to the legislation (2011), beginning with the 2013-2014 school year, school districts must ensure that the performance evaluation system and observation tool implemented at the local district level meet the advised criteria from the State of Michigan (Michigan Department of Education, 2011). The MCEE recommended that the common statewide tool be monitored to ensure proper local implementation to ensure reliability and fairness. Continued efforts at the state and local level are important to ensure the new evaluation system is implemented with fidelity and to address any factors that impact implementation within districts across Michigan.

The MCEE (2013) recommended the state provide sufficient base funding, as well as training for principals to support local school districts with the use of the state-selected tool. With increased accountability and demands in regard to teacher evaluation, principals will be responsible for implementation at the local level. With the high-stakes nature of the evaluation system, it is important that school districts and principals receive ongoing support, training, and resources to help guide the process of implementation at the local district level (MCEE, 2013).

This quantitative study aimed to determine if years of service, level of school, position, prior teaching experience, evaluation training, and gender influenced principal perceptions and the implementation of the newly defined teacher evaluation mandates in the state of Michigan.

Problem Statement

The new mandates within the state of Michigan require school districts to adopt and implement a rigorous, transparent, and fair performance evaluation system to evaluate all teachers annually while providing timely and constructive feedback. The Michigan Department

of Education (2012) reported that during the first two years of implementation districts were allowed to implement or maintain their own unique systems that met the basic requirements of the legislation. Districts across Michigan identified over 800 unique evaluation systems in place, with no system in place to check that instruments and procedures were implemented with integrity and rigor. The MDE reported that due to the diversity of evaluation systems across the state, direct comparisons of district effectiveness ratings and systems were quite difficult (Keesler & Howe, 2012).

The goal of the new system of evaluation is to enhance instruction, improve student achievement, and support ongoing professional learning in a more consistent manner across local districts. The new process will require a continued effort to ensure the evaluation system is being implemented with fidelity and the goals of the system are being met: to enhance instruction, improve student achievement, and to support ongoing professional learning. Without consistent implementation, the intended outcome will be hindered (MCEE, 2012).

Implementation research highlights the complex nature of the implementation process, with a continued focus to identify factors that drive successful program implementation (Blau, 2010; Honig, 2006; Fixsen et al., 2005; Fixsen et al., 2009). Implementation research suggests that the ways in which local actors, in this instance the school principal, make sense of and use such policies to determine the nature of the changes that actually occur in schools (Halverson, Kelley & Kimball, 2004; Desimone, 2002; Spillance, Reiser, & Reimer, 2002). With increased accountability at the local level, principals, as building level leadership, will play a critical role in the implementation process of the new mandates. In light of the recent reform and legislative measures in Michigan, it is necessary to provide a more in-depth look at the perceptions and experiences of principals in regard to the newly defined teacher evaluation mandates, due to their

role in the implementation process, to determine if these factors affect implementation at the local level.

As districts transition toward a more rigorous evaluation system, increased responsibility is placed on principals, with additional weight placed on their judgments in regard to educator practice (Graham, Milanowski, & Miller, 2012). Evaluation results are tied to high-stake decisions related to tenure, continued employment, placement, and compensation. With such decisions tied to evaluation results, it is important that state policies ensure school principals are highly- trained and certified to conduct the new teacher evaluations to increase fairness and consistency in the implementation process (Graham, Milanowski, Miller, 2012).

This quantitative study surveyed principals to assess their experiences and perceptions in response to the new teacher evaluation mandates to determine if the defined variables impacted local implementation in public schools across Michigan. The following section will define the purpose of the study and the research questions that guided the study.

Purpose of the Study

The purpose of this study was to determine whether principals' prior experiences influenced their perceptions in regard to the newly defined teacher evaluation mandates and therefore the implementation of the mandates within their local school district. This study accomplished its purpose by evaluating the variables to determine if principals' years of service, position (principal or vice-principal), type of school, prior teaching experience, evaluation training, gender, and perceptions influenced implementation of Michigan's new teacher evaluation mandates at the local district level. Principals are required to lead the process and are responsible for proper local implementation.

11

Due to recent reform and increased accountability measures across the nation and within the state of Michigan, it is important for stakeholders to remain aware of continued reform efforts and the outcomes of the efforts. The dramatic reforms involving teacher evaluation across the United States have led to promising and potential trends, but it is clear that policy is only part of the process because even the best evaluation system can be implemented poorly or undermined (National Council of Teacher Quality, 2012). Michigan has relied on teacher evaluation research, states with history of successful evaluation systems, and educational experts in the development of the new teacher evaluation system within the state. Currently, Michigan is only in the initial stages of implementation of its new teacher evaluation system. Because individual districts had the ability to select their own evaluation systems that met the state criteria, variations were identified within the system. Final recommendations from the MCEE were presented in the fall of 2013, and the success of the new evaluation will depend, to a large extent, upon the evaluation tools chosen by the districts and the fidelity of their implementation. The implementation efforts will be led, in many instances, by building principals.

This study sought to determine whether principals' prior experiences influenced their perceptions in regard to the newly defined teacher evaluation mandates and the implementation of the preliminary evaluation process within their local school district. The following will outline the research questions that guided the study.

Research Questions

In this study the researcher sought to determine whether principals' years of service, position, type of school, prior teaching experience, evaluation training, and gender influenced their perceptions in regard to the new teacher evaluation mandates and whether these variables impacted the implementation of the mandates at the local school district level.

1. To what extent does a principal's years of service, position (principal or vice-principal), type of school, prior teaching experience, evaluation training, and gender influence their perceptions in regard to the new teacher evaluation mandates in the state of Michigan?

2. To what extent does a principal's years of service, position (principal or vice-principal), type of school, perceptions, prior teaching experience, evaluation training, and gender impact the implementation of formal teacher observation and informal teacher observation in response to the new teacher evaluation mandates in the state of Michigan?

These research questions guided the study regarding the implementation of the new mandates in the process of formal and informal teacher evaluation. The conceptual framework that guided this study examined the factors that impacted implementation of Michigan's new teacher evaluation mandates as defined in The Implementation Drivers Framework (Michigan Department of Education, 2011). The Implementation Drivers include core components or drivers identified through research that are conducive to effective program implementation in the educational setting. This research study sought to identify whether the identified variables impacted implementation of the new teacher evaluation mandates at the local district level.

Significance of the Study

Teacher evaluation is a policy target in the United States to increase accountability and improve student learning and achievement, and Michigan has responded with legislation and a change in law to increase accountability in regard to teacher evaluation at the district level. With continued research and increased priorities, teacher evaluation has undergone reform measures and the new legislation within the State of Michigan required districts and principals to adapt to the new mandates.

Educational research is a continuous process and the new mandates within Michigan will warrant continued efforts by the Michigan Department of Education, the MCEE, educational experts, and local districts to ensure the evaluation system is implemented with fidelity and rigor,

and most importantly to ensure that the system is meeting the intended goals of enhancing instruction, improving student achievement, and supporting professional learning.

Over 18 months, the MCEE developed recommendations for a statewide teacher evaluation system. Specific research related to the new mandates is not yet available due to the initial stages of implementation. Michigan lawmakers and the MCEE are focused on reform to ensure a more valid, rigorous, and reliable teacher evaluation process. In large part, principals are responsible for making sense of the new mandates and responsible for proper implementation at the local level, highlighting the need to develop a deeper understanding of their experience in the process of implementation and the factors of influence.

Definition of Key Terms

The following definitions are provided to increase understanding of key terms used throughout the proposal.

A Blueprint for Reform - The 2010 reauthorization of the Elementary and Secondary Education Act. The Blueprint for Reform re-envisioned the federal role in education with a focus on key priorities: college-and-career ready students, great teachers and leaders in every school, equity and opportunity for all students, raising the bar and rewarding excellence, and promoting innovation and continuous improvement (United States Department of Education, 2010).

Classroom Observation- Used to measure observable classroom processes, including specific teacher practices, holistic aspects of instruction, and interactions between teachers and students. (National Comprehensive Center for Teacher Quality, 2011).

Formal Evaluation - Feedback as part of an evaluative process that is designed to assist professionals to perform to higher levels of mastery (Fisicaro, 2000).

Formative Evaluation – An evaluation process with the purpose of enhancing the professional skills of teachers (Danielson & McGreal, 2000).

Formal Observation- A formal observation is scheduled and includes pre- and post-observation conferences and written feedback and/or comments (United Federation of Teachers, 2012).

Highly Qualified - To be deemed highly-qualified a teacher must have: 1) a bachelor's degree, 2) full certification or licensure, and 3) prove that they know each subject they teach (Ed.gov, 2005).

Implementation - To carry out; put into action; perform (Dictionary.com, 2012). A mission-oriented process involving multiple decisions, actions, and corrections (Metz, Blase, & Bowie, 2007).

Informal Observation- Brief, unannounced visit that can occur at any time during the school day or class period, with the sole purpose of helping teachers develop and grow (Zepeda, 2012).

No Child Left Behind - The 2002 reauthorization of the Elementary and Secondary Education Act (ESEA). No Child Left Behind increased federal pressure on states to pursue standards-based reform agenda for public schools (United States Department of Education PL 107-110, 2012).

Race to the Top - Race to the Top is a competitive reform initiative that rewards states that are leading the way with ambitious plans for implementing coherent, compelling, and comprehensive reform in education (United States Department of Education, 2011).

Summative Evaluation – An evaluation process with the purpose of making consequential decisions regarding employment (Danielson & McGreal, 2000).

Teacher Evaluation - The process by which teachers are observed in a school setting and provided with feedback that is reflective of their performance (Fisicaro, 2000).

Tenure - A system of due process and employment guarantees for teachers (EdSource, 2009).

Limitations of the Study

This study was limited to principals and vice-principals within the state of Michigan and was administered at the end of the 2012-2013 school year. The timing and schedule constraints at the end of the school year may have impacted the response rate of individuals. The researcher also assumed that those that responded had more interest in the topic of teacher evaluation. The participants that did not respond to the survey may have a different perspective that will not be identified because of non-participation.

Summary

The new teacher evaluation mandates in Michigan require school districts to adopt and implement a rigorous, transparent, and fair performance evaluation system to evaluate teachers annually while providing timely and constructive feedback. Due to increased accountability at the local level, it is imperative to identify factors that impact implementation of the new teacher evaluation system. This study aimed to assess principals' experiences and perceptions in response to the new teacher evaluation mandates to determine if the defined variables impacted local implementation in public schools across Michigan.

Chapter two will provide the conceptual framework, a review of related literature in regard to teacher evaluation and the reform movement in the United States and the impact at the state and local level.

CHAPTER II

REVIEW OF LITERATURE

Overview

Teacher evaluation has been a focus of reform for many years, with the most recent reform measures impacting state and local decisions in regard to teacher evaluation. Increased accountability measures have influenced evaluation research to dig deeper to identify the most reliable, rigorous, and consistent method for a more comprehensive teacher evaluation process. Evaluation research continues to identify the need for improved evaluation systems, with a focus on key elements that support effective evaluation systems, measures for preparing and supporting principals to conduct evaluations, and an increased concentration on factors that influence implementation (Grossman, 2011).

Michigan, currently in the initial stage of implementation, worked to develop a more rigorous evaluation system for teachers for use in public schools across the state. The Michigan Council for Educator Effectiveness (MCEE) worked for 18 months to develop final recommendations for a statewide evaluation system based on evaluation research and the MCEE design principles. The new mandates within Michigan require school districts to adopt and implement a rigorous, transparent, and fair performance evaluation system to evaluate teachers annually.

With increased accountability, it is important to focus on the factors that influence implementation to ensure the evaluation system is meeting the intended goals and to inform future reform measures. The implementation process is a complex mission-oriented process "involving multiple decisions, actions, and corrections" (Metz et al., 2007). Principals have and

will continue to play a key role in the multifaceted implementation process with the new evaluation mandates.

As districts transition toward a more rigorous evaluation system, principals will be required to evaluate all teachers every year. Evaluation results will be tied to high-stakes decisions related to teacher tenure, continued employment and placement, and compensation. With decisions tied to evaluation results, it is imperative that principals are provided adequate resources, training and support to ensure consistent implementation at the local level (Graham, et al., 2012). It is also imperative that the evaluation process is monitored to ensure continued fidelity, with a focus on the factors/drivers that influence implementation.

This literature review focuses on the historical framework of teacher evaluation and the progression of reform measures that have impacted and continue to impact educational decisions in the United States and in Michigan. This chapter includes the conceptual framework, a historical background, evaluation research, and existing literature to frame the research study with a focus on principals' perceptions and prior experiences to examine their influence on policy implementation. By examining principals' prior experiences and perceptions in regard to the new teacher evaluation mandates in Michigan, this research study sought to develop a deeper understanding of the factors of influence within the implementation process.

Conceptual Framework

The conceptual framework that guided this study was developed using the Implementation Drivers framework and the MCEE teacher evaluation system design principles (MCEE, 2012; Michigan Department of Education, 2011). The Implementation Drivers framework identifies factors that influence successful policy and program implementation in the in the educational sector to include competency, capacity, and leadership drivers (Michigan

Department of Education, 2011). The MCEE teacher evaluation system design principles were developed from state law to guide the development of the statewide educator evaluation system. The final recommendations for a statewide educator evaluation system are integrated to ensure districts are provided with the competency and capacity to implement the new evaluation system. This study examined the competency, capacity, and leadership drivers involved in the implementation process of the new teacher evaluation based on the MCEE design principles.

Implementation research seeks to understand why a program operates and perform as it does and to identify elements of effective program implementation. Often new programs or policies are initiated by executive or legislative mandates, and local agencies are left to implement the new programs or to make changes in their existing programs. When mandated changes are extensive, this leads to the creation of new programs. Local agencies are often required to implement the changes or new programs within their environment. The implementation process is complex and begins with the initial stage of implementation prior to full operation, innovation, and sustainability (Fixsen et al., 2005).

One main focus of the implementation process is to ensure the local agency has the resources and capacity available and in place to get the program up and running, and to promote continued implementation with fidelity. It is necessary to provide local agencies with the materials, tools, equipment, procedures, and adequate time to prepare for program implementation (Michigan Department of Education, 2011; Werner, 2004). Getting the program up and running is only the beginning, with continued efforts throughout the process to ensure the program is suited to the environment, operating as planned, and most importantly is meeting the intended outcomes. Implementation research addresses these factors and seeks to develop a deeper understanding of the implementation process based on individual agencies (Werner,

2004). With amplified accountability and policy mandates, the educational sector continues to experience increased burden for local policy implementation. Increased accountability in the educational sector has created a need for research focused on the implementation process that goes beyond the initial stages of implementation (Blau, 2010; Honig, 2006; Fixsen et. al, 2005; Fixsen et al., 2009).

Following the initial stages of implementation are the stages of full operation, innovation, and sustainability. Full implementations of best practices occur once the new learning becomes integrated into practitioner and organization practices, policies and procedures. Following full implementation, the innovation stage of implementation is the stage of opportunity for local agencies to refine and expand the practice. Experts recommend that organizations implement the program with fidelity before attempting to innovate, so it is clear the agency is not attempting to adapt the practice to avoid the innovation. After full-implementation the local agency is responsible for the long-term sustainability and continued effectiveness of the implementation site in the context of a changing environment (Fixsen et. al, 2005). Once a practice is implemented within the educational sector, the local agency is responsible for continued sustainability often with decreased supports in place. Practices are frequently discarded at this stage of implementation with little effort to identify the reason for failure. Implementation research has reexamined the implementation process with a focus on the educational sector to identify drivers that are conducive to effective implementation (Fixsen et. al, 2005; Fixsen et al., 2009).

The Implementation Drivers framework was adapted from prior implementation research conducted in relation to policy and evidence-based practice implementation in the human services sector (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005). This research synthesis

review focused not only on the evidence-based programs, but also identified effective and efficient methods to translate policy mandates into effective implementation action. In a follow up to the synthesis review, Fixsen, Blase, Duda, Naoom, and VanDyke (2009) identified implementation drivers within the public school sector. Implementation drivers were defined as, "methods to develop, improve, and sustain component uses of innovations" (Fixsen et al., 2009, p. 13) with the goal of effective implementation in the educational environment.

The Implementation Drivers framework includes the following core implementation drivers identified by implementation research: competency drivers, capacity drivers, and leadership drivers. "The combination of the implementation drivers works to develop and engineer a 'host environment' that is conducive to the successful implementation of the program/initiative" (Michigan Department of Education, 2011, p. 1). Competency drivers are mechanisms to develop, improve, and sustain one's ability in the implementation process. The drivers include selection, training, and coaching. The capacity drivers are mechanisms to help create organizational environments that make it easier in the implementation process, including resources, information, and decision support data systems. Leadership drivers include vision, management/coordination, and facilitative administration to provide direction for the practices and allocation of resources. Together these drivers provide a host environment conducive to the successful implementation of a program, such as the new teacher evaluation system mandated by Michigan law.

Within Michigan, the MCEE developed recommendations for a statewide educator evaluation system to provide local districts with the capacity and competency for effective implementation. In the development of a statewide evaluation system, the MCEE outlined a clear set of design principles to reference in the process of creating a more rigorous and transparent

evaluation system. It is important to note that the MCEE principles are based on the new law and teacher evaluation research, and at this time, the evaluation system is not fully implemented. Instead, the MCEE has recently completed the pilot study and is analyzing the results of the study with continued efforts to ensure the recommendations meet the intended goals of the improved system. These principles include the following requirements: expectations should be clear and rigorous, the system should involve multiple measures, and the system should enhance performance. These design principles guided the MCEE in the development of the final recommendation for a statewide teacher evaluation system.

The design principles established by the MCEE were integrated with each guiding principle essential in the development and implementation of a statewide educator evaluation system. These design principles provided the MCEE with a clear outline in the development process and are evident in the final recommendation of a statewide evaluation system. The final recommendations include an evaluation system with clear and rigorous expectations, multiple measures to evaluate teacher effectiveness, and measures to enhance teacher performance. The statewide evaluation system intends to provide districts with the resources, information and training to promote a more consistent implementation process. Resources, information and training are defined within the implementation framework as necessary drivers in the implementation process. The implementation drivers are both integrated and compensatory as well, with each driver connected to the others aimed to improve local district's implementation effectiveness (Michigan Department of Education, 2011). The combination of competency, capacity, and leadership drivers develop a host environment conducive to effective implementation. The design principles and implementation drivers are both essential to ensure

proper implementation of a statewide evaluation aimed to improve the teacher evaluation process within the state of Michigan.

Although this study was guided by the implementation drivers framework, the questions within the study specifically examined the competency and capacity drivers involved in the implementation process in regard to the new teacher evaluation mandates within Michigan. The study examined the competency driver to include training and information provided to districts for implementation of the new mandates, as well as the capacity driver to include the resources provided to districts in the process of implementation. The questions within the study specifically focused on the amount of hours a principal received formal and informal training in regard to the new evaluation mandates and the expectations for the new process of evaluation and a principal's perceived understanding of the criteria and rubric language.

Principals, in large part, will be responsible for implementation at the local level, so this study focused on their role in the process of implementation and the factors that influenced the process. The following section will outline the implementation drivers and the role of the building principal and vice-principal in the implementation process.

Leadership

The Leadership drivers are mechanisms to provide direction for the practices and allocation of resources in the implementation process. Principals in Michigan will be responsible for providing the vision and coordination to successfully guide the implementation process within their local districts (Michigan Department of Education, 2011). The implementation skill of leadership involves a leader that is able to carry out programs and plans to successful completion. Building principals will be responsible for managing and facilitating the new evaluation system throughout the stages of implementation.

In order to manage, facilitate and coordinate the new evaluation system, principals must first develop an understanding of the requirements and vision defined by the MCEE. Policy implementation is often a complex process due to the educational environment and individual response to policy (Blau, 2010). "With changes occurring all over the world through different spectrums of understanding, it is inevitable that every person will experience change in different ways" (p. 1). In response to change and reform people often respond by attempting to make sense of the situation based on their own prior experiences and perceptions to guide their decision-making.

Once principals have developed an understanding of the new system they will be responsible for managing and facilitating the implementation process within their district. Principals will first need to define the desired outcome of the evaluation system and communicate the vision to the staff. Principals are responsible for the management and coordination of the evaluation system to ensure continued successful implementation and to review the progress of the implementation to make adjustments as needed.

The new mandates require principals to manage and facilitate an evaluation system that requires a shift in the process, as well as a shift in their responsibility as principal. The new mandates require principals to implement a more rigorous and transparent teacher evaluation system, with all teachers evaluated annually. "Institutionalizing educator evaluation for every teacher in every school multiple times across the year will require major changes in the work of the school principal" (MCEE Interim Progress Report, 2012, p. 4). This shift will require a strong commitment from districts and most importantly principals who will be responsible for implementation at the local level. Principals have an essential role in the entire teacher

evaluation process as a vehicle to improve teaching and learning with the potential to enhance the quality of teaching and student learning (Davis, Ellet, & Annuziata, 2002).

Competency

Implementation research identified this component as "competency drivers", which focus on, "developing individuals' knowledge and skills to implement the practices/program with fidelity" (Michigan Department of Education, 2011, p. 4). Providing leadership with adequate training, information, and resources to develop individuals' knowledge and skills are important mechanisms that help develop and improve one's ability in the implementation process (Fixsen et. al, 2005; Fixsen et al., 2009; Metz, Blase, & Bowie, 2007). The MCEE identified training and retraining of principals as a key element in a successful evaluation system.

Within the first two years of implementation of the new evaluation law in Michigan, principals received little or no concerted training in conducting evaluations, and this was identified as an issue to be recognized when evaluating reported teacher effectiveness ratings (Keesler & Howe, 2012). The MCEE outlined the importance of providing principals and leaders with adequate support and training and resources to ensure successful implementation with the new teacher evaluation system. "Using untrained evaluators significantly threatens the integrity and fidelity of the implementation, which in turn compromises both its capacity to improve student learning as well as its validity and reliability" (MCEE, 2012, p. 6).

Providing principals with training and retraining are important mechanisms that help develop, improve, and sustain the principal's ability in the implementation process (State of Michigan, 2011). The capacity drivers are also identified mechanisms to help create and sustain environments conducive to implementation.

Capacity

Within the Implementation Drivers, the capacity driver is defined as the process of, "engineering the organizational host environment to enhance the individual's performance" (p. 3). In relation to this study, the host environment is the local school and the individual performance is that of the principal in the process of implementation. The drivers include a data system for organizing and analyzing data, information to clearly and accurately explain the desired outcome of the evaluation system, and necessary resources provided for local implementation.

Within the capacity driver, examples are included for providing guidance on how to effectively implement the program with fidelity, data systems to support decisions, appropriate resources to perform all tasks, and continued feedback to ensure implementation is meeting the intended outcome of the program. The MCEE identified the importance of providing districts with the necessary resources, support, and materials for proper implementation as well. The resources will include clear and rigorous expectations to guide implementation, an observation tool, a student growth assessment, data systems for organizing and analyzing data, and continued guidance, training, and feedback to ensure implementation with fidelity.

Due to the high-stakes nature of the decisions that will be made based on these observations, "it is imperative that there be a system in place to check that instruments and procedures are implemented with integrity and rigor" (MCEE, p. 7). The MCEE stated, "doing nothing less would jeopardize the integrity of the entire process, limit the policy's capacity to improve schooling for Michigan's children, and compromise the entire reason for this initiative" (p. 7). The MCEE worked to develop recommendations for a fair, transparent, and feasible teacher evaluation system to ensure a more rigorous and consistent approach across the state of

Michigan. The new teacher evaluation system will require a strong commitment from local districts and building principals to ensure consistent and proper implementation.

The new process will require a continued effort by the MDE, the MCEE and local districts to ensure the system is being implemented with fidelity. Education is a complex arena in which policy implementation varies based on the organization, "in such contentious, interconnected, and multidimensional arenas, no one policy gets implemented or is successful everywhere all the time" (Honig, 2006, p. 2). It is important to focus on not only what is implementable and what works, but also focus on the conditions that lead to successful policy implementation in an organization (Honig, 2006). With increased accountability, it is important to focus on the factors that influence implementation to ensure the evaluation system is meeting the intended goals and to inform future reform measures. The Implementation Drivers guided this quantitative study in the process to examine factors to determine if these factors influenced the implementation process.

With increased accountability at the state level, local districts will be responsible for implementation of the new teacher evaluation mandates. The mandates have resulted in an ongoing process at the state level, with concentrated efforts by the MCEE to establish a more reliable and rigorous evaluation system. Evaluation research is an ongoing topic of inquiry, with increased accountability and the high-stakes nature of evaluation, past research and the history of evaluation reform are important to consider in the progression of new reform.

Previous Literature

In today's culture an increased priority has been placed on the teacher evaluation process in order to enhance instruction and student achievement. Various studies have been conducted in regard to past teacher evaluation systems with a consensus that many teacher evaluation models

were ineffective with little impact on the quality of teaching and learning in schools (Danielson & McGreal, 2000; The Education Trust-Midwest, 2010). Teacher evaluation has been a topic of reform throughout history in the United States and continues to evolve in response to reform measures based on the defined educational needs in our nation.

Historical Review of Teacher Evaluation

Throughout history, teacher evaluation has undergone numerous reform measures, with continued reform today.   Many evaluation systems used in districts were considered outdated and failed to meet the needs of 21st century learning environments. Many systems of teacher evaluations relied heavily on the documentation of a small number of observable behaviors, such as writing the learned objectives on the board or smiling at students (Danielson & McGreal, 2000). These are observable behaviors that can be checked off on an evaluation checklist during a single classroom observation. With increased accountability due to national and state educational reform measures, a checklist evaluation is not sufficient to meet the needs of the modern teacher evaluation goals (Danielson & McGreal, 2000; The Education Trust-Midwest, 2010: The New Teacher Project, 2010). Modern evaluation research recommends that evaluators conduct multiple observations per year with a research-based observation tool to provide a more reliable estimate of instructional quality (MCEE, 2013).

The teacher evaluation process has undergone numerous changes and these changes are often reflective of the defined educational needs determined within our society. This section of the literature will examine a historical review of teacher evaluation and reform in the United States.

In the early 1700s, education was not considered a professional discipline and local government and clergy were given the power to hire and fire teachers and make judgments about

their teaching (Marzano, Frontier & Livingston, 2011). Clergy were often given this role because of their extensive education and presumed ability to guide religious instruction in schools and teachers were simply considered servants to the community. At this time there was no agreement on specific criteria or pedagogy to guide the evaluation of teaching and the quality and type of feedback was highly diverse for all teachers. Yet supervisors of instruction were given this responsibility and the power to establish criteria to define instruction with the capability to hire and fire teachers (Marzano, et al., 2011).

During the 1800s, public education was seen as an essential component of the new democracy, to ensure all students were prepared as citizens of the United States. The common school movement pushed citizenship and academics to provide all individuals opportunity through education. The common school movement also created a demand for teachers who held expertise in specific disciplines and administrators who could assume the increasingly complex administrative role. Clergy were not trained in the specific subject areas and were no longer considered to have the expertise in making informed judgments about teacher effectiveness (Tracy, 1995). Often superintendents traveled from school to school, to encourage more effective instructional practices. The view of teaching had shifted, and teaching was viewed as a complex endeavor that required feedback in order to foster teacher expertise (Marzano, et al., 2011).

During this time supervision began to focus on improving instruction and pedagogical skills were considered to be necessary components of effective teaching, but there was no formal discussion about the specifics of these skills (Marzano et al., 2011). Teacher evaluation existed, but the main purpose of evaluation was to determine job continuation and pay increases. The

standards of the evaluation were often decided at the local level, with standards based on local education objectives (Markley, 2004).

Following World War II, education began to swing away from the scientific approach to schooling. Rather than focus on school as a factory model, teachers began to be treated as individuals. Evaluation emphasized the teacher as an individual and the process of evaluation to help develop his or her unique skills with consideration of his or her emotional needs. In the 1940s and until the 1950s teacher quality was judged from a moralistic and ethical perspective (Ellett & Teddlie, 2003). Educators and researchers emphasized presage variables, which were identified as teacher traits: voice, appearance, emotional stability, trustworthiness, warmth, and enthusiasm. During this era it was believed that if a teacher possessed these traits they were more likely to perform effectively, so these traits became a main component in the checklist-type teacher evaluation (Danielson & McGreal, 2000).

During this era supervisors experienced increased responsibility in the school environment. Supervisors were required to evaluate teaching based on classroom visitations with a focus on the teacher's management of the physical environment. Principals were also responsible for supervision of curriculum and resources for instruction, as well as meeting with teachers on an individual basis. Although teacher evaluation still had room for improvement, a positive outcome of this era was the consensus of the importance and utility of teacher observation in the classroom (Marzano et al., 2011). Including classroom observation led to further reform efforts to reflect evaluation that promoted growth.

The clinical supervision movement, developed in the late 1950s, spread rapidly in the field of education (Marzano et al., 2011). The clinical supervision model, used in teaching hospitals, focused on the process of a purposeful and symbiotic relationship between the

practitioner and resident with observation and discussion between both parties to encourage growth and effectiveness (Goldhammer, 1969). The model was then refined to work with student teachers and practicing teachers. The process included specific steps in the method of supervision: pre-observation conference, classroom observation, analysis, a supervision conference, and a follow up analysis to encourage teachers and supervisors to engage in a reflective process to promote growth and continuous improvement. The process intended to encourage the interaction of the observer and teacher, with the phases aimed to be guiding steps to help identify and encourage effective instructional practices. The clinical supervision model did not always result in rich dialogue; rather evaluators viewed the process as steps to follow in the evaluation procedure. Although evaluation reform progressed, this model was still widely used in many schools throughout the 1980s (Marzano et al., 2011).

In the 1960s and 1970s a shift occurred within our country to increase student performance in order to compete globally, especially in math and science. The United States began pushing for basic skills acquisition and to improve science and math instruction. Kersten and Israel (2005) described this time in school where evaluation and curriculum became a focus of reform. Curriculum was developed around a factory model with leveled readers and independent learning models in math and science. The teacher's role was to guide students and encourage progression in the curriculum with a continued focus on traits such as voice, warmth, and enthusiasm. The building principal's role as the manager was to ensure this progression occurred and ensure teachers were keeping accurate records to demonstrate student progress.

An increased spotlight on teacher evaluation research occurred during this same time period, and many researchers focused on teacher behaviors in the classroom to determine if any behaviors were linked to student achievement (Danielson & McGreal, 2000). This research was

referred to as teacher effectiveness research and later named the research on teacher effects. The research attempted to determine if connections did exist among teacher behavior and student achievement, and in some cases it demonstrated a connection between teacher behavior and basic skill acquisition. Danielson and McGreal (2000) stated, "it was important work in that it began to form the basis for a set of fundamental teaching skills that are a part of the current framework for teaching" (p. 13). Danielson and McGreal (2000) attempted to identify what teachers did or could do to improve student basic skills.

School reform efforts of the 1980s created an increased attentiveness to teacher evaluation as a critical lever for improving the quality of teaching (Darling-Hammond, 1990). "By 1983, 26 states required teacher evaluation in some form" (Daley & Kim, 2010). In the early 1980s Madeline Hunter and her colleagues at UCLA began to look at teaching from a theory-based perspective. Hunter developed a set of prescribed teaching practices designed to improve teacher decision making to enhance student learning. During this time many schools developed evaluations based on these prescribed teaching methods and behaviors. Evaluations included rating scales and checklists to evaluate teachers and these scales encouraged a single view of teaching. Principals were armed with checklists looking for the prescribed teaching model in each classroom. The Hunter model had its benefits, but had no consistent evidence to demonstrate positive effects on student learning. The teacher effect research and the Hunter model encouraged an emphasis on teacher-centered, structured classrooms, with increased instructionally focused staff development (Danielson & McGreal, 2000; Marzano et al., 2011).

The teacher effects research and the Hunter model influenced the way in which state policymakers and local school districts developed evaluation practices. State and local evaluations were often based on evaluation criteria generated from lists of behaviors from the

teacher effects research and the Hunter seven steps in lesson design. The rating scales and checklists created for implementation in the local districts encouraged a single view of teaching. "Despite the misinterpretation and misuse of the effects research, this approach made a significant contribution to education because it clearly confirmed the critical role that teachers play in student learning" (Danielson & McGreal, 2000, p. 14).

Research and reform focused on teacher evaluation continued into the late 1980s and 1990s once again with pressure to increase student-learning outcomes. The United States began focusing on the U.S. economy and the changing demands in the workforce to prepare students for the changing job market. The societal pressures as well as teacher effect research contributed to a noticeable shift in the study of teaching once again (Danielson & McGreal, 2000). Due to the increased pressure on student performance with a push for more complex outcomes such as critical thinking and problem solving, the definition of good teaching was reexamined. "Research on the interdependence of teaching methods and content has promoted new understanding about the dynamic interplay among content, teacher, learner, and context that must always be accommodated if teaching is to be effective" (Danielson & McGreal, 2000, p. 14). In order to evaluate these expanded views of teaching and more complex student outcomes, more reliable and valid forms of student assessment were necessary. This also resulted in a demand for a more reliable and valid forms of teacher evaluation.

In the late 1980s, Danielson (1986) published work based on supervision and evaluation to provide a research- based framework for teaching, seeking to capture the dynamic process of classroom teaching (Marzano et al., 2011). This model was considered the most detailed and comprehensive approach to evaluation at the time. The Danielson Model included four domains: Planning and Preparation, the Classroom Environment, Instruction, and Professional

33

Responsibilities. Within each of these domains, Danielson outlined components required to demonstrate the knowledge, skills, and dispositions required in demonstrating competency in classroom teaching (Marzano et al., 2011).

During classroom observation, principals looked for these components to rate the teacher based on their instruction. The framework included 76 elements of quality teaching, all broken into four levels of performance: unsatisfactory, basic, proficient, and distinguished. The intent of the framework was to honor the complexity of teaching, constitute a language for professional conversation, and provide a structure for self-assessment and reflection of professional practice (Danielson, 1996). Although this model was identified as the most detailed and comprehensive approach to evaluation at the time, the framework has been updated based on modern research to include the same four domains, with an increased focus on teacher growth.

The historical overview of teacher evaluation lead to some distinct characteristics in evaluation reform that are also evident in today's reform efforts. Throughout history the teacher evaluation process shifted to reflect modern research findings within each era. As educational research progressed, the purpose of evaluation was more tightly defined with an increased focus on improving teaching and promoting growth. The same conversations are occurring in today's reform efforts. Continued research efforts have defined the purpose of evaluation once again with attention to align evaluation systems to drive instructional improvement through evidence-based feedback based on multiple measures.

Past reform measures focused on many strategies for improving school development, without enough attention paid to teacher evaluations (Toch, 2008). "Teacher evaluations are at the very center of the education enterprise and can be catalysts for teacher and school improvement" (Toch, 2008, p. 32). This is evident with the increased amount of educational

research focused on identifying criteria for effective teacher evaluation systems that recognize highly effective and ineffective teachers (Danielson & McGreal, 2000; Marx, 2007; Toch, 2008). The following section will highlight the continued need for teacher evaluation reform, the importance of identifying effective and ineffective teachers, recommendations for evaluation system reform, and continued efforts at the national and state level to increase accountability and promote growth through teacher evaluation.

Teacher Evaluation as a Target for Reform

Educational legislation and reform measures continue to increase demands and accountability in public education across the United States with heightened attention on teacher evaluation as a means for improving teacher performance and student achievement. Traditional teacher evaluations have failed to provide teachers with accurate evaluation ratings or information to improve their practice (Sartain, Stoellnga, Brown, Luppescu, Matsko, Miller, Durwood, Jlang, & Glazer, 2011). Teachers are the single most significant factor in-school predictor of achievement, which demonstrates the need for an evaluation process that has high expectations and is committed to teacher development (Center for American Progress, 2011; MCEE, 2012).

Many states across the U.S. are working to develop comprehensive evaluation systems that are more rigorous and reliable with a concentrated focus on teacher development and student achievement. Evaluation research has provided states and local school districts with specific recommendations in the development of teacher evaluation systems that evaluate teachers fairly and accurately. Research recommends the following components in the development of a more comprehensive evaluation process to include:

Clear and Rigorous Expectations

Multiple Measures

Measures to Support Ongoing Educator Learning and Development

(Daley & Kim, 2010; Goe, Bell, & Little, 2008; Little, Goe & Bell, 2009; MET, 2013; The New Teacher Project, 2010; Weisberg, Sexton, Mulhern, & Keeling, 2009).

The following section will describe the key components recommended through education research in the development of comprehensive evaluation systems committed to teacher development.

Comprehensive Teacher Evaluation

The high stakes nature of teacher evaluation has led to increased accountability and decisions based on evaluation outcomes, thus creating the need for increased accuracy within evaluation frameworks. Well-designed teacher evaluation systems can have a direct and lasting effect on individual teacher performance (Taylor & Tyler, 2012). Several initiatives and research partnerships have resulted in the development of practical guides and guidelines for reference in the process of creating more rigorous teacher evaluation systems to evaluate teacher effectiveness in a more consistent manner. Evaluation research has outlined key elements necessary for inclusion in the development of more rigorous teacher evaluation systems aimed to drive instructional improvement to include clear and rigorous expectations, the inclusion of multiple measures, and measures to support educator learning and development.

Clear and Rigorous Expectations

Clear and rigorous expectations are necessary to guide the development of new evaluation systems, but are also necessary to provide continued clarity and rigor for principals

and teachers in the entire implementation process. Teacher evaluation systems often fail to

evaluate teachers in a rigorous way due to the lack of an accurate system to measure teacher

effectiveness (Goe et al., 2008). Teacher evaluations should include precise language with a clear

and coherent definition of exemplary practice with little room for inferences to ensure that both

principals and teachers interpret the expectations in the same way (Danielson, 2012). In the

process of evaluation reform, systems must first develop definitions of effective teaching and

clear performance standards, multiple rating options with unambiguous expectations within each

category, as well as specific protocol for implementation throughout the entire evaluation

process.

The definition of effective teaching and methods for measuring teacher effectiveness

have evolved and changed over the course of education due to changing demands in

accountability and changes in the defined purpose of evaluation (Goe et al., 2008). With an

increased focus on accountability and student growth, teacher effectiveness and measures for

evaluating effectiveness in a consistent manner are a main focus in evaluation reform. Effective

teachers are believed to be the most critical component in the classroom and the most important

school-related factor in student achievement gains (Daley & Kim, 2010). Although research has

highlighted the importance of identifying effective and ineffective teachers in the classroom,

performance is seldom evaluated in a rigorous way with evaluations failing to differentiate

between the teachers who boost student learning and those that need to improve (Daley & Kim,

2010; Sartain, Stoellnga, Brown, Luppescu, Matsko, Miller, Durwood, Jlang, & Glazer, 2011;

Weisberg et al., 2009). Our nation's schools have failed to distinguish great teaching from good,

good from fair, and fair from poor (Sartain et al., 2011; Weisberg et al., 2009). Without accurate

systems to measure and identify teacher effectiveness and areas for improvement the process

fails to impact teaching and growth. The following section will highlight studies that have examined teacher effectiveness ratings and the inconsistency in the process of evaluation.

The Widget Effect (2009) examined the nation's failure to recognize and respond to the variations in the effectiveness of teachers due to the lack of a teacher evaluation system to assess these variations. School districts assumed classroom effectiveness was the same from teacher to teacher, treating teachers as interchangeable parts rather than individual professionals. The study was conducted within 12 districts across four states, with strikingly similar outcomes. In districts that used two ratings (generally satisfactory and unsatisfactory), more than 99 percent of teachers received the satisfactory rating. Districts using a broader range of ratings reported 94 percent of teachers in the top two ratings, with less than one percent rated as unsatisfactory. When nearly all teachers are rated in the highest two categories, excellence goes unrecognized and poor performance goes unaddressed. Failure to recognize variations in instructional performance also prevents districts from identifying needs in their teachers. Teacher evaluation systems espouse grand intentions for development, assessment, and improvement, but evaluations rarely achieve these goals (Weisberg, et al., 2009, p.19).

Another study concluded similar findings within their traditional teacher evaluation system. In 2008, the Chicago Public Schools (CPS) launched the Excellence in Teaching Pilot, aimed to revamp the teacher evaluation process (Sartain et al., 2011). The study included data from the traditional evaluation system for comparison to the new pilot evaluation system implemented in the CPS. With the traditional evaluation system principals used an observation checklist to rate teacher performance in a number of areas with ratings: strength, weakness, or does not apply, with no definition to define strength or weakness. At the end of the year, principals provided teachers with a final performance rating. According to the traditional

38

system, 93 percent of the teachers rated in the top two rating categories with only 0.3 percent identified as unsatisfactory, at the same time that 66% of the Chicago Public Schools were failing to meet state standards (Sartain et al., 2011). The traditional system failed to differentiate among teachers with varying skill sets, with so many teachers receiving the highest ratings the best teachers were not rewarded and the lowest were neither supported nor sanctioned. Both principals and teachers reported the checklist system to be neither meaningful nor fair.

Teacher evaluation research has indicated that teacher evaluation ratings fail to differentiate among teachers with varying skill sets and fail to provide accurate ratings for teacher effectiveness. (Sartain et al., 2011, Weisberg, et al., 2009). Another study examined data to compare effectiveness ratings of individual teachers and concluded inconsistent ratings when ratings were compared from year to year (Newton, Hammond, Haertel, & Thomas, 2010). Ratings data from five school districts were inconsistent when examined from class to class and from year to year. Teachers who scored in the bottom 20% rankings in one year, only 20% to 30% had similar ratings the next year, with 25% to 45% of the teachers ranked at the top of the distribution, scoring well above average (Newton et al., 2010). Similar results were found for those who scored at the top of the ratings, a small minority stayed in the same rating band the following year. Traditional evaluation systems fail to accurately rate teacher effectiveness and fall short in providing teachers with appropriate feedback to promote growth.

With teacher evaluation as a target for reform, research has focused on identifying measures to improve rater reliability within evaluation frameworks. The CPS (2011) study focused on identifying ways to increase observation score reliability. When teachers were given the same total number of observations and included the perspective of two or more observers per teacher, reliability was enhanced greatly. The CPS pilot study measured reliability of ratings by

comparing principal and observer ratings of the same lesson. Most principal ratings were consistent with the ratings of external observers when it came to rating teaching practice, though 11 percent of principals consistently gave lower ratings than the observers and 17 percent consistently gave higher ratings than the observers (Sartain et al., 2011). The Measures of Teacher Effectiveness (2013) study reported that adding lessons and observers increased the reliability of classroom observations ratings as well. If a teacher's results were based on two lessons split between multiple observers, the reliability was greater than that achieved by a single observer (MET, 2013). Additional observers, including observers from outside a teacher's school were found to increase reliability by providing added perspectives and a continuous check for in-school bias (MET, 2013; Sartain et al., 2011).

Research has indicated inconsistencies in the ratings of teachers in the evaluation process, with continued efforts to identify means to improve rater reliability. Studies have found that reliability increased when multiple ratings were included, a clear definition of teacher effectiveness was provided, multiple lessons were observed, and addition of a second observer (MET, 2013; Sartain et al., 2011). Clarifying the way teacher effectiveness is defined is important for teachers, principals, and policy makers, because teacher effectiveness can be measured with the appropriate measurement tool and process (MET, 2012). The definition of teacher effectiveness should be a broad and comprehensive definition to truly capture the complex responsibilities of a teacher and to guide the efforts in developing appropriate teacher effectiveness measurements. In response to research literature, policy, discussions, standards, and reports a comprehensive definition of teacher effectiveness was conceptualized (Goe et al., 2008):

- Effective teachers have high expectations for all students and help students learn, as measured by value-added or other test-based growth measures, or by alternative measures.

- Effective teachers contribute to positive academic, attitudinal, and social outcomes for students such as regular attendance, on-time promotion to the next grade, on-time graduation, self-efficacy, and cooperative behavior.

- Effective teachers use diverse resources to plan and structure engaging learning opportunities; monitor student progress formatively, adapting instruction as needed; and evaluate learning using multiple sources of evidence.

- Effective teachers contribute to the development of classrooms and schools that value diversity and civic-mindedness.

- Effective teachers collaborate with other teachers, administrators, parents, and education professionals to ensure student success, particularly the success of students with special needs and those at high risk for failure. (p.8)

It is important for all stakeholders to have a clear understanding of the attributes of an effective teacher, including principals and teachers. In order for principals to access teacher effectiveness, they must have a clear understanding of the attributes of an effective teacher. Teachers need clearly defined performance standards to work toward and principals need a common view of quality teaching with concise tools to consistently evaluate teachers (Danielson & McGreal 2010; New Teacher Project, 2010). Lack of consistency in teacher effectiveness ratings highlights the continued efforts for teacher evaluation systems to clearly define teacher effectiveness and clear performance standards.

Clear and rigorous expectations provide states and local districts with precise language and definitions, clear performance standards, and more consistent expectations across the board in the development and implementation of teacher evaluation systems. Traditional evaluation systems have failed to accurately identify effective and ineffective teachers and fail to drive instructional improvement. With increased accountability traditional evaluation systems are no longer adequate. More comprehensive evaluation systems to accurately evaluate teaching and drive instructional practice should incorporate clear and rigorous expectations with a focus on evidence-based feedback based on multiple measures of evaluation. The following section will focus on the research findings and recommendations for including multiple measures in the evaluation process.

Multiple Measures

Increased understanding in brain research has provided a more complex picture of teaching and learning and this complexity carries over into reform (Kersten & Israel, 2005). "Education is built around a conception of practice based on current and emergent research findings; as those findings suggest new approaches, pedagogical practices must also move forward" (Danielson, & McGreal, p.1). A shift will need to occur from the typical checklist evaluations to a more advanced evaluation system in order to meet the new criteria. The evaluation process should serve both as a formative and summative purpose, in providing summative evaluation scores and formative feedback to improve teachers' practice (Goe et al., 2008). Teacher evaluation systems must become more aligned with high expectations and drive instructional improvement through evidence-based feedback based on multiple measures.

A comprehensive evaluation of teacher effectiveness should be based on a composite of multiple measures, rather than a single measurement because any single measurement is prone to

error (MET, 2013). Multiples measures of evaluation that go beyond student achievement scores should be incorporated to capture a more comprehensive picture of teacher effectiveness (Sartain et al., 2011). Multiple measures can include: classroom observations, value-added models, analysis of classroom artifacts, portfolios, self-report of practice and evidence of teachers' contributions to the school as whole (Goe et al., 2008; Little et al., 2009; MET, 2012). Finding ways to combine multiple measures that support effective teaching and learning will be a challenge for local districts. The following section will outline multiple measures recommended by research for a more comprehensive approach to measuring student growth to include classroom observations, classroom artifacts, portfolios, and value-added models.

Classroom Observations

Classroom observations are a key part of the evaluation process, but research indicated the process of observation was not as effective as it could be. Classroom observations require a time commitment from teachers and principals and in order to justify this time commitment, classroom observations should be accurate, reflective, and reliable (Ho & Kane, 2013). Modern research has identified several steps in order to increase accuracy and reliability of classroom observations to include the implementation of specific and validated observation protocol and more frequent and informal observations (Goe et al., 2008; MCEE, 2012).

Recent research has indicated the importance of specific and validated observation protocol for the evaluation process in order to increase the reliability of the outcomes (Goe et al., 2008; Ho & Kane, 2013; MET Project, 2013). When states or districts choose an instrument for implementation, careful attention should be paid to its intended and validated use (Goe et al., 2008). Principals should be fully trained to develop a clear understanding of the protocol and

use of the tool with continued efforts throughout the implementation process to assure the protocol is implemented with fidelity.

Part of the evaluation protocol should include more frequent observations of teachers throughout the school year with all teachers evaluated annually to provide constructive feedback aimed to drive improvement (Goe et al., 2008; Little, et al., 2009; MET, 2012; New Teacher Project, 2010). "Annual evaluation is the only way to ensure that all teachers-regardless of their ability levels or years of experience-get the ongoing feedback on their performance that all professionals deserve" (New Teacher Project, 2010, p. 4). Within the traditional evaluation system, a principal collected all evidence during one classroom observation. One observation is not enough and simply conducting walkthroughs is not sufficient in providing an in-depth evaluation (MET, 2013).

Taylor & Tyler (2012) focused on one approach to teacher evaluation: practice-based assessments, which rely on multiple highly- structured observations conducted by experienced peer teachers and administrators. This study compared the achievement of a teacher's students during the year that he/she was evaluated to the achievement of the same teacher's students in the years before the evaluation year. The study found suggestive evidence that the effectiveness of individual teachers improved during the school year in which they were evaluated, with a continued increase in the years after evaluation. "We estimate that the average teacher's student score 0.11 standard deviations higher in years after the teacher has undergone an evaluation compared to how her students scored in the years before her evaluation" (Taylor & Tyler, 2012, p. 83). The findings within this study suggest that teachers develop skills and learn new information about their performance that encourage a change in a lasting manner as a result of

the subjective performance evaluation through multiple observations conducted by experienced teachers and administrators.

More frequent formal and informal observations of teaching are important in the process of evaluation to provide evaluators with a true representation of a teacher's practice. Formal classroom observations are scheduled and include pre- and post- observation conferences with written feedback and comments from the evaluator. Formal observations can include the use of evaluation rubrics, teacher portfolios, observation checklists, and the development of teacher professional growth plans. Formal classroom observations are no longer the only type of observation conducted within the evaluation process.

In recent years, informal classroom observations have evolved in literature and in practice with the continued urgency to improve the teacher evaluation process. The informal observations can be tied to the management by wandering around method in the business sector. This method encourages the leader to make unannounced visits to employees, encouraging informal conversation and communication to provide opportunity to watch, observe, and listen to learn more about the employee on a personal level. These types of observations are encouraged in the educational sector as well, tied to both summative and formative evaluation practices. An informal observation is defined as a brief, unannounced visit that can occur at any time during the school day or class period, with the sole purpose of helping teachers develop and grow. Informal observations can include classroom walkthroughs, learning walks, or informal conversations (Zepeda, 2012). Informal observations provide principals with an opportunity to build relationships with staff, motivate teachers, and to keep informed about classroom instructional practices (Zepeda, 2012).

Frequent classroom observations, both formal and informal provide principals with more opportunity to collect data to accurately measure teacher effectiveness. More rigorous evaluations require various measures in the evaluation process other than the traditional teacher classroom observation. Evaluation systems should include multiple measures in the process of evaluation to increase reliability.

Student Growth

Increased accountability has led to an amplified focus on student achievement and proficiency in learning. Teaching is defined as effective when it enables student learning, but identifying effective teaching in relation to student learning is a complex matter, due to the diversity of students and their achievement levels (Danielson & McGreal, 2000; MET, 2013). Teachers are the most important school-based determinant of student achievement causing a growing interest in teacher evaluation systems that link teacher and student achievement (Goe et al., 2008; MET, 2012). In response to increased accountability measures, more effective approaches for assessing teacher effectiveness in relation to student growth have been sought.

Teacher effectiveness is often regarded in reference to a teacher's effectiveness based only on student achievement measured by standardized tests, but other evidence of student growth should be also be included (Goe et al., 2008). The development of assessments to measure student achievement is a continued effort at the national and state level to ensure accuracy and reliability in measuring student growth. "Approaches to teacher evaluation that incorporate a measure of student learning require valid techniques to assess that learning" (Danielson & McGreal, 2000, p. 41). Validity concerns the extent to which evidence demonstrates that an assessment measures what it is intended to measure and provides sound evidence for decision-making purposes (Herman, Heritage, & Goldschmidt, 2011). In the

educational sector, measurements should be validated prior to use as a student growth measure.

Valid measures could include evidence found in the classroom such as school-level assessments, district assessments, pre and post-test measures, projects, and other measures to identify student growth. An Arizona study (2012) focused on a program that required teachers to use various methods of student assessment to compliment evaluations of their practice. The study found that over time teachers improved their ability to create tools to assess student learning, development of pre and post tests, and greater awareness of the importance of sound curriculum development with a focus on high quality content and instructional strategies (Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012). Continued efforts are necessary to provide principals and teachers with support to develop valid measures for inclusion in the evaluation process.

Classroom artifacts such as student work and projects are another measure to compliment evaluation of practice. Analysis of artifacts can provide summative information about instruction and rich formative information (Goe et al., 2008). If evaluators ask teachers to select student work to provide concrete evidence of student achievement, teachers are almost guaranteed to learn from the reflective process (Danielson & McGreal, 2000). This process requires teachers to take an active role in the evaluation process and requires teachers to ask questions, dig deeper, and reflect on student learning in the classroom. Teachers who have an active role in the evaluation process enhance their teaching performance, resulting in enhanced student performance (Packard & Dereshiwsky, 1991). Although the process has shown to encourage an active role of teachers, more research is necessary in order to validate the use of this method as a primary measure in evaluation (Goe et al., 2008).

Increased interest in identifying student growth has also led to the development of value-added models to measure teacher effectiveness based on student performance. There are many approaches to value-added models, ranging in complexity, but value-added models share a common goal: to identify the effect of an individual teacher on student learning. Value-added models have proved to be controversial and research efforts continue to work to identify and develop reliable systems to accurately measure the effect of an individual teacher on student learning.

The MET project (2013) sought to determine if measures of teaching could identify teachers who help students learn more. The project first estimated teachers' effectiveness using multiple measures for the 2009-2010 school year, then principals created rosters for each class and the rosters were assigned randomly within each grade and subject, the project predicted student outcomes based on teachers' previous results, and then compared those predictions to actual differences. In order to account for the diversity of the student population, the teachers within the study were randomly assigned to students for two consecutive years. Following the random assignments, the adjusted measures identified teachers who produced higher and lower average student achievement gains.

> On average, students of teachers with higher teacher effectiveness estimates
> outperformed students of teachers with lower teacher effectiveness estimates. Moreover,
> the magnitude of students actual gains largely corresponded with gains predicted by their
> effectiveness measured the previous year. (MET, 2013, p. 8)

The research confirmed that teachers previously identified as more effective caused students to learn more than those teachers with lower effectiveness estimates.

The Chicago Public Schools (CPS) also studied the relationship between teaching practice and student learning to test the validity of observation ratings within their system. Each teacher was rated based on classroom observation ratings, measured by the Danielson Framework, and given a value-added score. To test validity, a statistical model was used to compare the Framework ratings to the value-added measures. Across almost all of the Framework components, teachers with the lowest observation ratings also had the lowest value-added measures, and the value-added measure increased when a teacher's rating increased. Students showed the most growth in classroom with highly rated teachers and students showed the least growth in classrooms with teachers with low observation ratings. The study identified a "strong relationship" between classroom observations ratings (measured by the Danielson Framework ratings) and value-added measures (p. 9). This pattern generally held across the components in math and reading, suggesting the Framework was measuring teaching practice in a valid way (Sartain et al., 2012, p. 11).

Holtzapple (2003) conducted a similar study in the Cincinnati schools using the Danielson Framework for Teaching (2006) to compare scores using a value-added model of predicted achievement versus actual achievement. Similar findings were reported within this study, the author found correlation between the observation scores and the value-added scores for teachers. "There were significant and positive correlations between the teachers' total rating and the mean gains for their students" (Holtzapple, 2003, p. 215). Teachers who received unsatisfactory and basic ratings on the Teaching for Learning Domain had students who performed lower than predicted on state and district tests, based on prior year test scores. Teachers with advanced or distinguished ratings generally had students that performed better

than expected, and teachers rated proficient had students that generally performed close to the expected level.

Heneman, Kimball, and Milanowski (2006) conducted a multiyear mixed-methods study investigating teacher evaluation systems and value-added models in four sites in the United States, including Cincinnati. The evaluation tool examined was a modification of Danielson's Framework for Teaching (2006) and a value-added model in which achievement was estimated based on prior achievement and student characteristics. The study reported positive relationships were found between teacher evaluation scores and student achievement gains, but there was substantial variability across the sites and within the sites. The findings were not consistent across sites, suggesting that using value-added strategies instead of classroom observation does not necessarily result in more valid assessment (Heneman et al., 2006).

Heneman et al. (2006) speculated that the links found between teachers performance and student test scores was in part dependent on the performance of the classroom evaluators, not just the performance of the teachers. Holtzapple (2003) reported high correlations between observations and value-added scores within the study in Cincinnati, but Heneman et al. speculated that the high correlations were due to the highly trained raters conducting the evaluations.

Value-added models remain controversial in evaluation reform and require ongoing research to evaluate their use in a more comprehensive evaluation system. The complex process of measuring and linking student growth to teacher evaluation has sparked disagreement in educational reform. Harris (2011) stated that, "value-added refers to statistical techniques that make predictions based on information that includes but goes beyond prior student achievement" (p. 2). The predictions are made and then compared with what actually happens. Schools that

50

have students that reach higher achievement levels than predicted have high valued-added scores, and schools with students that end up with scores lower than predicted receive low value-added scores. Harris argued that value-added models focus too highly on high-stakes testing which result in teaching to the test, narrowing of curriculum, and in some instances cheating. "Value-added measures have potential, but we can't lose sight of the larger purpose: measuring performance in a way that facilitates genuine accountability, a sense of mission, and sound messages" (Harris, 2011, p. 5). Value-added models are a continued focus in evaluation research, with efforts to establish more valid and reliable ways to assess teacher effectiveness and student growth.

Teacher evaluation systems must become more aligned with high expectations to drive instructional improvement through evidence-based feedback based on multiple measures. Multiple measures of evaluation will provide a more comprehensive picture of teacher effectiveness to provide individual teachers with opportunity for growth and development.

Support Ongoing Educator Learning and Development

The evaluation process should be a means to make reasonable judgments regarding the quality of teaching, a means for identifying teachers in need of support, and aimed to drive instructional improvement (Danielson & McGreal, 2000). Teacher evaluation systems "should give teachers feedback on classroom needs, the opportunity to learn new teaching techniques, and counsel from principals and other teachers on how to make changes in their classrooms" (Boyd, 1989, p. 1). Yet most evaluation systems are based on a one-way type of top-down communication, with the only evidence of teacher performance collected by the administrator during classroom observations (Danielson & McGreal, 2000).

The typical observation is conducted by a principal or supervisor during a single classroom observation, with the principal taking notes followed by a write up of the observation with performance feedback based on the single classroom observation. "In the absence of clear evaluative criteria, this feedback is likely to be highly idiosyncratic, and may or may not be of value to the teacher" (Danielson & McGreal, 2000, p. 5). A single classroom visit is not adequate to collect enough data to identify effective and ineffective teachers to provide evidence-based feedback aimed to improve instruction.

Evaluations fail to recognize excellent teachers and chronically low performing teachers with little differentiated support to guide teachers to improve as professionals (New Teacher Project, 2012). Within the New Teacher Project study, teachers reported rarely receiving meaningful constructive feedback on their performance about how to improve their practice through the formal evaluation system. Evaluators were found to spend no more time observing or providing feedback to teachers identified as mediocre or poor performers than with highly rated teachers. The study found that teachers who received lower ratings reported the same number of observations and the same amount of feedback as reported by their highly rated colleagues. Feedback can be a mechanism for principals and teachers to drive improvement, but without a system in place to accurately rate teachers, the process of providing feedback is ignored.

There are many different reasons for measuring teacher effectiveness, but a key reason is to identify weaknesses in instruction and to collect information in the process that is useful in designing professional growth plans to improve instruction (Goe et al., 2008). Traditional evaluation systems fail to assess variations in instructional effectiveness, which prevents principals from identifying areas of need for teachers. The Widget Effect study reported 73% of

teachers surveyed said their most recent evaluation did not identify any development areas, and only 45% of teachers who did have development areas identified received useful support to improve (Weisberg et al., 2009, p.4). Teacher evaluation should be linked to professional development based on teacher need. Evaluation results should be analyzed to make decisions about professional learning needs at the teacher, school, and district level (Goe, Biggers, & Croft, 2012).

An increased priority should be placed on guiding teachers in the process to improve and principals play a major role in that process. Ball (2012) discussed the importance of focusing on teacher evaluation in a new way with an increased focus on the feedback process. "So the kind of evaluation that we need to be building is one in which the kind of observations that are conducted of teaching really zoom in on the actual work and allow teachers to improve" (Ball, 2012). Learning is a continuous process and teacher evaluation should also be continuous.

Evaluation should promote teacher growth and effectiveness with support differentiated based on teacher need (Danielson & McGreal, 2000; Weisberg et al., 2009). Evaluation research indicates that new evaluation systems should incorporate more frequent classroom observations, follow up meetings to provide feedback, and increased focus on identifying teacher need based on evaluation results. (Little, et al., 2009; MET, 2013; New Teacher Project, 2012).

Teacher evaluation continues to be a topic of reform with continued efforts to identify the most reliable and rigorous method of teacher evaluation. Well-designed teacher evaluation systems can have a direct and lasting effect on individual teacher performance when teacher effectiveness is accurately identified and teachers are provided feedback and support based on individual need.

With increased accountability measures across the United States, research efforts, both public and privately funded, have resulted in an influx in the development of research initiatives and partnerships aimed to improve teacher evaluation. Schmoker (2012) advised the educational sector to proceed with caution with the influx of complex teacher evaluations due to their premature implementation and lack of solid evidence that these evaluations do in fact promote better teaching. With such an increase in accountability and reform measures, it is important for local agencies to proceed with caution when developing and adopting a teacher evaluation system to ensure the new system meets the intended outcomes.

Increased accountability has led to the development of research partnerships, ongoing research efforts, and national and state policy measures aimed at development of a more comprehensive evaluation systems that drives instructional improvement through evidence-based feedback. Based on evaluation research, the key elements in a more comprehensive evaluation system include clear and rigorous expectations, incorporation of multiple measures, and ongoing efforts to encourage teacher development. In response to national accountability measures, states across the U.S. are working to develop more comprehensive evaluation systems. The State of Michigan is currently in the process of developing a more comprehensive evaluation system and the following section highlights legislation and continued efforts at the state level in creation of a new statewide teacher evaluation system. The following section will also go into depth in regard to implementation at the local level and the factors that have been found to impact policy implementation.

<div align="center">Teacher Evaluation as a Target for Reform in Michigan</div>

In response to national reform measures Michigan passed teacher evaluation legislation that required local districts to implement a more rigorous, transparent, and fair teacher evaluation

process to include observation tools and evaluations that are examined for reliability and validity. Developing a more rigorous evaluation system was necessary in Michigan, due to the lack of a consistent evaluation system to meet the demands and rigor of national and state reform efforts. The Michigan Council for Educator Effectiveness was established and appointed to develop common tools and expectations for statewide implementation.

Teacher Evaluation in Michigan

Hazi and Rucinski (2009) reported on the analysis of state statutes and department of education regulations in fifty states to determine the affect on the practice of teacher evaluation and the implication for instructional supervision resulting from the policy actions since the passage of No Child Left Behind in 2002. Hazi and Rucinski reviewed teacher evaluation as a policy target for improved student learning across the fifty-states to determine the impact of legislation at the state level on the practice of teacher evaluation. No Child Left Behind (2002) and the National Governors Association (NGA) targeted teacher evaluation as a means to instructional improvement to improve student learning by ensuring a highly qualified teacher in every classroom. The NGA (2002) identified six policy goals for improving student learning:

    1. Define teacher quality.

    2. Focus evaluation policy on improving teacher practices.

    3. Incorporate student learning into teacher evaluation.

    4. Create professional accountability through developing career ladders.

    5. Train evaluators in pre-service programs and broaden participation in evaluation

    designs.

The purpose of the study was to determine the extent in which the NGA goals appeared in individual state statutes and regulations and to determine how the changes affected the practice

of teacher evaluation within the fifty states. The following section highlights the findings of the study pertaining to Michigan.

At the time of Hazi and Rucinski's (2009) study, each state determined the level of state control over teacher evaluation practices. The state of Michigan was defined as a level 1 control, meaning the least prescriptive in which the state department delegated choice and control of the evaluation policy, criteria, and instrument to be determined by the local school district. The states defined as level 2 allowed the local districts to determine the evaluation policy, criteria, and instrument with government approval. The states defined as Level 3 were more involved locally by specifying the criteria as to the process of teacher evaluation. At level 4 the state specified the instrument and procedures of teacher evaluation. The various levels provided for great diversity of teacher evaluation expectations across the United States. With Michigan defined as a level 1, local school districts had the control over determining the evaluation policy, criteria, and the instrument used in the teacher evaluation process. This created a great deal of inconsistency in the United States, as well as across school districts in Michigan.

Hazi and Rucinski's (2009) study also reported the state departments that adopted strategies for oversight for analysis of the teacher evaluation process within the state. Michigan reported no strategy adopted for oversight at the time of the study. Within Michigan the evaluation process, criteria, and choice of instrument in teacher evaluation were largely determined and defined by the local school districts. This could be problematic for local school districts that had no definite and consistent process in place for teacher evaluation to define and recognize teacher effectiveness and performance. Another problematic area concluded within the study was the lack of training provided for evaluators across the fifty states. The study indicated that many states were beginning to implement training programs for evaluators, but

Michigan had not yet adopted a training strategy for evaluators. Hazi and Rucinski's report indicated Michigan's need for a more comprehensive and consistent evaluation system across the state even after the passage of No Child Left Behind mandates.

To increase the nation's investment in education once again, the federal government introduced the competitive grant program, Race to the Top. The Race to the Top (2010) grant was designed to encourage and reward states that demonstrated educational innovation and reform plans within their state with a focus on increasing student achievement. Several criteria were outlined within the grant program to identify states that have demonstrated innovation and educational reform. One of the criteria for selection focused on ensuring that all schools have great teachers and leaders, with a concentrated effort on successful methods for improving teachers and principal effectiveness.

Although Michigan was not awarded Race to the Top funds in the second round, the State of Michigan passed educational reform legislation with a focus on the development of a more rigorous teacher evaluation system aimed to encourage teacher growth and student learning (Senate Fiscal Agency, 2010). The following section will highlight continued teacher evaluation reform within the state of Michigan and the work of the Michigan Council for Educator Effectiveness in their efforts in development of an improved teacher evaluation system.

In response to national reform measures the State of Michigan passed teacher evaluation legislation with continued efforts to develop a more rigorous system and efforts to ensure successful implementation at the local district level. Due to the increase in teacher evaluation accountability, the Michigan Council for Educator Effectiveness was established and appointed to develop common tools and expectations for statewide implementation. The MCEE was established with council members appointed in September 2012, with five voting members. Per

executive order (2012) The Governor's Council on Educator Effectiveness, created by Public Act No. 102 (2011) was renamed The Michigan Council for Educator Effectiveness (MCEE).

The MCEE was created as a two-year temporary agency to develop a "fair, transparent, and feasible" evaluation system for teachers and administrators based on rigorous standards of professional practice and measurement (MCEE Interim Progress Report, 2012, p. 1). Following the 2012-2013 school year and the pilot study, the MCEE submitted a report (2013) to the State Board of Education, the Governor, and the state legislature to recommend a new teacher evaluation system that included a student growth and assessment tool, a state evaluation tool for teachers and administrators, and a process for evaluating and approving locally developed evaluation tools.

The MCEE (2012) had a large responsibility in their work to develop and recommend a statewide evaluation system to improve the current teacher evaluation system with the goals of enhancing instruction, improving student achievement, and supporting ongoing professional development. The MCEE collaborated, researched and worked with experts in the process of developing a fair and transparent evaluation system for teachers across Michigan. The work of the MCEE is important within this study because of their key role in the development of a statewide teacher evaluation system for implementation at the local level.

The following section includes an in-depth look at the work of the MCEE and the clear design principles outlined for the development of a fair, transparent, and feasible educator evaluation system based on evaluation research recommendations.

New Mandates within the State of Michigan

National and state legislation indicated the importance of a teacher performance assessment that provides increased feedback for helping individuals grow and improve. School

districts are required to adopt and implement a rigorous, transparent, and fair performance evaluation system to evaluate teachers annually while providing timely and constructive feedback. The MCEE (2012) identified key design principles essential in the of development of the state's educator evaluation system that included:

- Expectations should be clear and rigorous.

- The system should involve multiple measures.

- The system should enhance performance.

- The system should be committed to and structured to support ongoing educator learning and development. (p. 3)

The MCEE also indicated the performance evaluation must meet the following criteria:

Evaluate the teacher's job performance at least annually, use multiple rating categories that take into account student growth with timely and constructive feedback, establish clear approaches to measuring student growth while providing teachers with relevant data on student growth, and districts must use the evaluation to inform decisions regarding promotion, retention, and development of teachers. (p. 3)

The MCEE relied on research, reports, and other states' experiences in regard to teacher evaluation reform to guide the process in Michigan. The MCEE identified key principles and criteria that were essential in the development of a more rigorous state educator evaluation system. One of the key elements of the new system included the development of a common evaluation instrument.

Common Evaluation Instrument

A large responsibility of the MCEE was the development of criteria for the observation process and a common evaluation instrument for implementation in the local school districts. The MCEE recommended five criteria for reference in the review and selection process of observation procedures and observation instruments/tools to be implemented in Michigan school districts. The five criteria included:

1. The instrument should be aligned with relevant state and national standards for educators.

2. The instrument should be used both for describing practice and supporting ongoing educator learning/development.

3. The instruments should be accompanied by a rigorous and ongoing training program for evaluators.

4. Independent research on the reliability and the validity of the instruments should be available.

5. The demands of the process should be feasible (in terms of personnel, time, and financial cost. (p. 3)

The MCEE conducted a pilot study to determine a common evaluation tool. Districts were selected based on geographic locations, demographics, and size in order to represent the various types of schools in Michigan in the pilot process. The teacher observations tools were based on four models of teacher evaluation: 5 Dimensions of Teaching and Learning, Charlotte Danielson's Framework for Teaching, the Marzano Teacher Evaluation Model, and The Thoughtful Classroom. Following the pilot year, the MCEE (2013) reviewed preliminary findings and indicated that current data from the study suggested little significant difference

among the tools. Final results from the pilot were not yet reported. After a complete analysis of the findings the MCEE will review all recommendations and make adjustments based on the new data in order to provide a final recommendation for a common observation tool for implementation in local districts.

Clear and Rigorous Expectations

Michigan's previous evaluation system lacked a clear definition of effective teaching and standards for evaluating teachers (Education Trust-Midwest, 2012). The prior evaluation law allowed for great variation across the state on evaluating teaching and measurement of student growth. The new guidelines in Michigan require principals to evaluate all teachers annually with clear and rigorous expectations to guide the process. The MCEE established an evaluation system based on clear and rigorous expectations to encourage greater consistency in the evaluation process across the state and within school districts. Evaluation should be designed to promote clear purposes and teacher evaluation systems have experienced scrutiny for their lack of clear purpose and expectation (Danielson, 2010; Pallas, 2012).

Classroom Observations and Annual Evaluation

A major problem in Michigan's teacher evaluation is the lack of valuable, meaningful, and timely professional feedback for teachers to help improve their practice (Education Trust-Midwest, 2012). This is not only a disservice to thousands of teachers, but also thousands of students in our state (Education Trust-Midwest). The new evaluation requires principals to evaluate every teacher at least annually.

In Michigan, the MCEE outlined criteria for the observation process to include a more rigorous observation system; to evaluate all teachers annually, to include pre and post

conferences with teachers, extended and brief observations with time to review and analyze the observational data collected, and adequate time to conference with each teacher. "The intent of supervision is to improve teaching and to lend assistance to teachers as they move through their careers. To this end, supervision is a proactive, ongoing set of processes and procedures" (Zepeda, 2012, p.12). Based on the pilot study and research, the MCEE recommended at least three classroom observations for the most reliable estimate of a teacher's performance. Multiple classroom observations, both formal and informal, lead to higher levels of reliability in the evaluation process (Measures of Effective Teaching Project, 2012).

Evaluation research recommends improved evaluation systems to include more frequent formal and informal observations, focused feedback following observation, and the implementation of specific procedures with trained evaluators (Danielson & McGreal, 2000; Little, et al., 2009; MET, 2013; New Teacher Project, 2012; Zepeda, 2012). In response to evaluation research, the MCEE focused on development of clear evaluative criteria and instruments to assist principals in the evaluation process to provide a more in-depth analysis of teaching. Although observations of teaching may seem simple to carry out, "the council's research makes clear the need to be vigilant in demanding rigorous and accurate use of instruments that have also been field-tested, their reliability and implementation analyzed, and critically reviewed" (MCEE Interim Progress Report, 2012, p. 7).

The MCEE evaluation recommendations (2013) include observation tools and instruments researched and piloted prior to implementation. Current data from the pilot study suggested little significant difference among the tools: Charlotte Danielson's Framework for Teaching, Marzano Teaching Evaluation Model, The Thoughtful Classroom, and 5 Dimensions of Teaching and Learning. The MCEE recommended the state should choose one of the four

piloted observation tools for statewide implementation. If final results from the pilot provide further evidence that suggest one of the tools to be less reliable or practical then the state should make adjustments as needed (MCEE, 2013). The purpose of the instrument in the evaluation process is to help identify underperforming and high performing teachers to encourage continual improvement and support both teachers and principals' learning.

Along with more frequent formal and informal observations, teacher evaluation should include multiple measures to provide a more comprehensive picture of a teacher's classroom practice. With a more rigorous evaluation process across the United States and in Michigan, principals will be responsible for developing an understanding of the process, with large responsibility for implementation at the local level.

Multiple Measures

Teacher evaluation reform measures have focused on development of more rigorous evaluation systems to include multiple measures in the evaluation process, with many states in the process of identifying effective measures for inclusion in their systems. Student growth will become a main component to determine a teacher's effectiveness in the classroom in order to meet the criteria defined by the law. For the 2013-2014 school year, at least 25 percent of the annual year-end evaluation should be based on student growth and assessment data, for the 2014-2015 school year, at least 40 percent of the annual year-end evaluation should be based on student growth and assessment data, and for the 2015-2016 school year, 50 percent of the annual year-end evaluation should be based on student growth and assessment data (Pubic Act 102, 2011).

The MCEE has reviewed a range of assessments, clarified the meaning of student growth, identified plans for the future of Michigan assessments, and identified challenges in the process

of combining observation and student growth scores. "The central purpose of teaching is to help students learn, and student growth measures can provide valuable insights into teachers' effectiveness, particularly when coupled with other measures of teaching efficacy" (MCEE, 2012, p. 7). Evaluation ratings were analyzed following the initial year of implementation of the new evaluation system with student growth measures connected to evaluation. Following the initial year of implementation the State of Michigan reported districts that included higher percentages of student growth in their evaluations were more likely to use all four ratings more frequently (Keesler & Howe, 2012). Evaluation ratings expanded in most districts from two ratings (satisfactory to unsatisfactory) to four rating categories (highly effective, effective, minimally effective, and ineffective), with most differentiation occurring between effective and highly effective.

> As the percentage of the evaluation based on student growth increase, the number of teachers reported as "highly effective," "minimally effective," and "ineffective" increases. In short, presence of more student growth data seems to be related to more frequent utilization of all four evaluation-rating categories. The data suggests that as a statewide educator evaluation system is developed and implemented, with prescribed percentage of student growth, districts will be able to better differentiate and more fully utilize all categories of ratings more frequently. (Keesler & Howe, p. 10)

The MCEE (2013) final recommendations encourage the state to continue to develop or select assessments that are aligned to state-adopted content standards in both core and non-core content areas and to provide when possible value-added modeling scores for educators based on the state-provided assessments. It is important to continue to address this component in the teacher evaluation process in order to provide an effective approach for measuring student

growth and linking observations and student growth scores to teacher effectiveness.

The MCEE released information in October 2012 pertaining to the consultations with several national vendors in the process of adopting a student growth tool. The national vendors included AIR, Pearson, SAS, and Wisconsin VARC. All vendors indicated an interest to the MCEE in the results of the pilot study in order to analyze the results to adapt their approach in measuring student growth. Common tools for measuring student growth are not yet identified, but the value-added model has been identified as criteria for development of the student growth tool.

The MCEE provided other potential components that could be included to support observations, as well as other materials provided by teachers, principals, and parents. The components are as follows:

> Pre-observation conferences, post-observation conferences, summative evaluation conferences, teacher self-assessments, professional accountabilities, educator growth plans, locally developed assessments of student learning, structured review of student work, teacher artifacts using portfolio or evidence binder processes, feedback from students, parents, and/or other teachers using structured survey tools, teacher self-reflection and progress on professional growth goals (MCEE, 2012, p. 11).

The MCEE concluded in their process thus far that it is important to consider carefully how values produced from observational and student growth tools are combined to achieve a final evaluation score. Districts in Michigan will be required to utilize the assessment to measure student growth and report the outcome as a component of the teacher evaluation. At year-end effectiveness labels must be reported to the Michigan Department of Education based on the annual evaluation.

The MCEE (2013) recommended using three categories for rating teachers rather than the initial four categories provided in the interim report and preliminary stages of the development process. The MCEE had significant concerns about the lack of reliability and usefulness of using four categories because they found no empirical or theoretical evidence to support the use of four categories. The MCEE instead recommended using the following three rating categories when evaluating teachers: professional, provisional, and ineffective. The final evaluation rating would incorporate quality of practice based on observations and quality of student outcomes based on student growth data.

Multiple measures in the evaluation process allow teachers to provide evidence of student learning and their own growth, but continued support is necessary in the process as well. Enhancing teacher performance is a main goal of improved teacher evaluation systems, and continued support for principals and teachers is part of the process to encourage improvement. Principals will have a major role in the entire teacher evaluation process with an increased push for proper local implementation, thus requiring increased support measures for principals responsible for implementation.

<center>Implementation at the Local Level</center>

The entire process of teacher evaluation is under reform including the evaluation tools, the defined procedure to include the amount of time, format, feedback, and the role of the principal (Danielson & McGreal, 2000) and the new mandates require principals to take a more in-depth role in the process. With the high stakes nature of the new teacher evaluation system and process it is imperative that school districts and principals receive ongoing support, training, and resources to ensure proper local implementation. If the core implementation drivers are not considered in the development process of a fair, transparent, and feasible evaluation system the

new mandates alone will not ensure successful implementation at the local level. The new process will require a continued effort to ensure the system is being implemented with fidelity and the goals of the system are being met to enhance instruction, improve student achievement, and support ongoing professional learning.

The conceptual framework that guided this quantitative study was focused on the Implementation Drivers necessary for successful implementation. "The combination of the implementation drivers works to develop and engineer a "host environment" that is conducive to the successful implementation of the program/initiative" (Michigan Department of Education, 2011, p. 1). The MCEE reviewed research and collaborated with states that had undergone evaluation reform to identify drivers that promote successful implementation and drivers that pose challenges to implementation.

An effective evaluation system is not simple to carry out and local school districts need support for implementation including observation guidelines, high-quality training and professional development, and continued supports to ensure integrity in the implementation process. Practitioners and policymakers have spent a good deal of time comparing the effectiveness of one rubric over another, yet the fairness and meaningful evaluation relies far more than the merits of a particular tool. An observation rubric is a tool that can be used effectively of ineffectively.

> Reliability and validity are functions of the users of the tool, as well as the tool itself. The quality of implementation depends on the principal and observer buy-in and capacity, as well as the depth and quality of training and support they receive (Sartain et al., 2011, p. 41).

Full implementation will require continued support for evaluators through training and retraining, adequate resources, a data system for collecting and analyzing data, and strong leadership in the management and facilitation of the new teacher evaluation system (MCEE, 2012).

Training

With new teacher evaluation systems, principals are responsible for a more rigorous observation and feedback process. Principals are required to evaluate their staff and make judgments based on their observations, so it is important that they are adequately trained in order to make accurate, consistent judgments based on evidence. "Observers should be trained to observe carefully, attend rigorously to the key elements of instruction, to be thorough and accurate in their note taking and assessments, and responsible in the conclusions they draw from their observations" (MCEE Interim Progress Report, p. 3). Providing training for principals and leadership involved in teacher evaluation is crucial in the process of successful implementation (MET, 2013; State of Michigan, 2011).

Providing information that is timely and understandable to those responsible for implementation is a key mechanism to support an individual's competency to implement the evaluation system with fidelity (Michigan Department of Education, 2011). Training research has defined effective training workshops to consist of the following: presenting information (knowledge), providing demonstrations (live or taped) of the important aspects of the practice, and assuring opportunity to practice key skills in the training setting (Fixsen et al., 2005). Principals need opportunity to develop an understanding of any new policy mandate, and new learning that requires a more complex repertoire of skills is more difficult for trainees to learn

and master, and demands greater planning and precision on the part of the trainer (Joyce & Showers, 2002)

Ball (2012) discussed the importance of providing principals with training to make sure they are watching things that really matter in the classroom with carefully planned tools to guide the observation focused on providing feedback to improve teaching. During the phasing in stage, as well as throughout the implementation process, it is important that educators and evaluators are properly trained with time to acquaint themselves with the tools and "reorient themselves to the changes in their responsibilities that are required by the system" (MCEE, p. 6). The MCEE has included training as a necessary component for all evaluators to provide for consistency across districts in order to ensure reliability of the system as a whole.

Principals will be responsible for implementation at the local level and their role in the process is detrimental to the success of the system. Newly developed evaluation systems will provide districts with a common framework and protocol for evaluation, but other factors are important to consider in the implementation process as well.

Principal Perceptions of Teacher Evaluation

Implementation research suggests that policy implementation is a complex arena, in which local stakeholders influence the process based on their own past experiences (Desimone, 2002; Honig, 2006; Spillane, Reiser, Reimer, 2002). Newly developed evaluation systems will provide school leaders with common frameworks to serve as a basis for improving teaching and learning in schools, but providing a framework is only part of the process, with the practice of implementation as a key factor in the success of the system. Various factors can influence and in some situations restrain implementation, including individual perceptions (Blau, 2010).

Perception is defined as an individual's way of regarding, understanding, or interpreting something (Oxford Dictionary, 2013). With new policy mandates in regard to teacher evaluation, principals must develop an understanding and interpretation of the new requirements in order to implement the mandates at the local level. Past experiences with evaluation, as well as individual bias can affect an individual's interpretation and understanding of the new evaluation mandates. Bias is defined as an individual's cause to feel or show inclination or prejudice for or against someone or something (Oxford Dictionary, 2013).

Biases are so deeply engrained into societal thinking that it jades decisions and progression of realistic solutions to common problems (Blau, 2010). Individual bias was found to contribute to failed implementation, affecting the overall outcome of the policy. Principals begin the process of evaluation with their own personally perceived purpose and bias, creating subjectivity with the final outcome (Blau, 2010). In the implementation process, principals will play a key role, as facilitator in providing the vision and coordination to successfully guide the implementation process, highlighting the importance to develop a deeper understanding of their past experiences and perceptions of teacher evaluation systems.

Henry-Barton (2010) conducted a quantitative study in regard to principals' perceptions of teacher evaluation practices in an urban school district and sited key findings related to evaluation. Principals indicated using both summative and formative evaluation as effective evaluation approaches for non-tenured teachers. The study reported finding no statistical significance in use of evaluation approaches when analyzing data based on principals' positions or years of experience. Yet the perceived effectiveness of evaluation approaches for non-tenured teachers varied when analyzing data based on the position of the principal. Elementary principals identified a mix of formative and summative assessment to be most effective in

70

providing teachers with feedback for promoting growth, middle schools principals identified the post observation conference to be most effective, and high-school principals perceived the observation checklist to be most effective (Henry-Barton, 2010, p. 114).

Data were also analyzed based on years of experience and concluded a difference in perceptions. Principals with 1-10 years of experience, indicated that five out of six approaches, both formative and summative, to be more effective for non-tenured teachers, and principals with 11-15 years of experience identified the pre and post observation conference as the most effective means to provide non-tenure teachers with feedback for promoting growth. Principals with more experience might have increased familiarity in balancing the evaluation process and their increased accountability as principal or principals may see the process as perfunctory and choose to utilize only some elements of the process in order to meet evaluation deadlines (Henry- Barton, 2010). Principals often see traditional evaluation programs as administrative burdens and perfunctory (Halverson, Kelley, & Kimball, 2004). Principals also reported that summative evaluations are time intensive and prevent other opportunities for principles to work with teacher to improve instruction (Kersten & Isreal, 2004). These perceptions may then impact the implementation or carrying out of the policy mandates.

With increased rigor in the evaluation process, time commitment was another area of concern for principals. Without adequate time to support the process of evaluation, the process may revert back to a less rigorous approach, impacting the effectiveness of evaluation (Kersten & Israel, 2005). There was evidence to suggest principals with 6-10 years of experience spent more time on non-tenured teacher evaluation, but insufficient evidence to suggest a difference in time spent in regard to principals with 1-5 years of experience and those with 11 years or more experience. Fifty-one percent of principals surveyed were responsible for evaluation of 11 or

71

more teachers per year. With more teachers to supervise in addition to administrative duties, there is less time to devote to the overall process (Henry-Barton, 2010).

Another quantitative study (Fisicaro, 2010) explored principals' perceptions of teacher evaluation and the characteristics of evaluation with different findings when analyzing data based on principals' various attributes. The results demonstrated that different groups of principals did not indicate differences in perceptions and attitudes about evaluation across the groups were generally consistent. The findings were consistent among principals of different genders, in different school settings, and principals with varying levels of experience. The results of the study indicated that regardless of school setting, years of experience, or gender, principals across the state of New Jersey shared similar views pertaining to teacher evaluation (Fisicaro, 2010).

Principals will be responsible for developing an understanding and interpretation of the new teacher evaluation mandates in Michigan for implementation at the local level. Principals' past experience and personal perceptions of the new process could impact implementation at the local level, highlighting the importance in developing an understanding of their past experience and perceptions to determine if these variables impact local implementation.

Summary

Educational reform continues to influence federal, state, and local policy due to the nature of increased accountability in education. Teacher evaluation has and continues to remain a topic of reform across the United States and within the state of Michigan. It is important for policy makers to understand the history of reform efforts and educational research to guide the process of current reform. The MCEE worked to review literature, consulted with other states, and conducted pilot studies in the process of developing an improved teacher evaluation system.

It is imperative that the state, local districts, and evaluators remain focused on the implementation process with vigilant efforts to ensure that implementation at the local level is consistent and meeting the intended goals of the system.

Past research has focused on principals' experiences in education and their perceptions of teacher evaluation in the quest to identify the most effective teacher evaluation practices, with less research conducted in regard to the relationship of principal perceptions and past experiences and the influence or impact on implementation. Chapter 3 will outline the research design for the study, research questions, population and sample, and methods of analysis for the study conducted.

CHAPTER III

METHODOLOGY

Overview

This study sought to determine whether principals' years of service, position, type of

school, prior teaching experience, gender and evaluation training influenced their perceptions

and impacted the implementation of formal and informal teacher observations in response to the

new teacher evaluation mandates in Michigan. Chapter 3 includes the research design,

population and sample, research questions, and methods of analysis for the study.

Purpose

Principals will be responsible for proper local implementation of the newly defined

teacher evaluation process to increase reliability in the execution of the new system. This study

sought to determine whether principals' prior experiences influenced their perceptions in regard

to the newly defined teacher evaluation mandates and the implementation of the mandates within

their local school district.

Research Questions

1. To what extent does a principal's years of service, position (principal or vice-principal), type of school, prior teaching experience, evaluation training, and gender influence their perceptions in regard to the new teacher evaluation mandates in the state of Michigan?

Independent Variable- Years of service, position (principal or vice-principal), type of school, prior teaching experience, evaluation training, and gender
Dependent Variable- Perceptions

2. To what extent does a principal's years of service, position (principal or vice-principal), type of school, perceptions, prior teaching experience, evaluation training, and gender impact the implementation of formal teacher observation and informal teacher observation in response to the new teacher evaluation mandates in the state of Michigan?

Independent Variable- Years of service, position (principal or vice-principal), type of school, perceptions, prior teaching experience, evaluation training, and gender
Dependent Variable: Implementation of Teacher Evaluation

Research Design

The research study utilized a non-experimental quantitative approach. The main purpose of this non-experimental study was to focus on the variables including perceptions, position, level of school, years of service, prior teaching experience, training, and gender of principals to determine if these variables predict the outcome, defined in this study as the implementation of formal and informal teacher evaluation observation practices. A non-experimental approach was chosen as the research method because the research study focused on how the proposed cause and the proposed outcome relate to each other, without providing a specific treatment to one group and withholding it from another (Creswell, 2009). A cross-sectional approach was used to collect data, because data was be obtained at one point in time, but from participants with varying attributes (Field, 2009).

Population

The population of participants in the study was public school principals and vice principals at the elementary, middle, and high school level in the state of Michigan involved in implementation of the newly defined teacher evaluation mandates. For the purpose of this study elementary was defined as K-5, middle school was defined as 6th-8th, and high school was defined as 9th -12th grade.

Sample

The new teacher evaluation mandates in the state of Michigan require principals to properly implement the teacher evaluation process at the local district level to evaluate all

teachers every year and to report yearly effectiveness ratings to the state of Michigan registry, thus indicating their importance as the identified population of study.

According to the Michigan Department of Education and the Center for Educational Performance and Information (2012) public data set there were 3,843 public school principals and vice principals in the state of Michigan. The sample was male and female principals and vice principals in the state of Michigan at the elementary and secondary level. The survey was administered to 1200 public school principals within the state of Michigan, with 180 public school principals completing the survey.

The sample was chosen using a stratified random sampling to ensure each individual had the equal probability of being selected from the population to increase credibility of the study (Creswell, 2009). Creswell (2009) stated, "stratification means that specific characteristics of individuals (e.g., both females and males) are represented in the sample and the sample reflects the true proportion in the population of individuals with certain characteristics" (Creswell, p. 148). Within this study, the stratified random sampling was first chosen based on the level of school to ensure an equal representation of principals from the elementary, middle, and high school level. The next sampling was based on the position of the principal to include an even distribution of principals and vice-principals from different schools, random sampling without replacement.

The survey was administered to 1200 public school principals and vice-principals, 400 at the elementary level, 400 at the middle school level, and 400 at the high-school level. Within each of these samples, 200 principals and 200 vice principals were chosen to survey to ensure equal representation of principals based on their position. The survey was sent by e-mail with a

link to the Qualtrics survey in April of 2013 with a second follow-up e-mail sent as a reminder in May 2013 and a third reminder sent in the beginning of June 2013.

Data Collection

No appropriate instrument for this study was found in research or literature, so the researcher developed a survey instrument to measure principals' perceptions of the constructs based on the new teacher evaluation mandates in Michigan in consultation with dissertation committee members. The constructs were developed based on the Michigan Council for Educator Effectiveness design principles for the development of the state's educator evaluation system, evaluation research, and prior dissertations. All committee members were consulted in the process of establishing validity of the survey instrument as well.

In order to develop a survey to adequately correspond with the new mandates within Michigan, the researcher focused on the MCEE Interim Progress Report (2012) which outlined the background of the MCEE and the charge and vision of the council in the process of developing a fair, transparent, and feasible evaluation system for teachers in the state of Michigan. In the ongoing process of development of a new teacher evaluation protocol and tool the Michigan Council for Educator Effectiveness has focused heavily on teacher evaluation research. The research that was reviewed in chapter 2, as well as the MCEE Interim Progress Report (2012), suggested that teacher evaluation should include expectations that are clear and rigorous, a system that involves multiple measures, and a system that is committed to and structured to support ongoing educator learning and development. These are the three constructs by which the survey was developed. The researcher also referenced Barton and Fisicaro's dissertations focused on measuring principals' perceptions of teacher evaluation to help develop

the survey for the study and utilized Fisicaro's dissertation in particular (Barton, 2010, &

Fisicaro 2010).

One of the studies that the researcher referenced in the process of survey development

was a dissertation based on principal perceptions of teacher evaluation. Fisicaro developed a 20-

question survey instrument to measure perceptions based on modern evaluation research and an

expert panel to devise valid survey questions. He selected the constructs from the review of

prior research and then developed questions to measure the various aspects of each construct.

The panel critiqued the survey based on the content of the survey and the support of research and

agreed "that the four constructs were valid and the questions chosen as a measure of each

construct were reasonable" (Fisicaro, 2010, p. 68). Fisicaro (2010) suggested the theoretical

constructs within the questionnaire were based on the theoretical background from research on

teacher evaluation. The four constructs incorporated were:

- Teacher evaluation should be formed in a partnership.

- Teacher evaluation should be ongoing.

- Teacher evaluation should consider student learning.

- Teacher evaluation should be differentiated for all individuals.

Fisicaro carefully devised survey questions based on research, as well as established an expert

panel of individuals to guide in the construction of the questions. He then conducted a pilot

study to determine any further revisions, which resulted in a question deletion prior to survey

administration. The constructs and survey questions from the Fisicaro study were reviewed and

analyzed to determine if the questions were aligned with the new mandates within the state of

Michigan. Fisicaro included teacher evaluation research by several of the same authors as

referenced in the MCEE research. The researcher incorporated ten questions from the Fisicaro

study in this study's survey for administration in Michigan. Table 3.1 indicates the questions
that were incorporated from the Fisicaro study.

Development of Questions

After selecting the constructs from the review of research and the MCEE design
principles, the researcher consulted with committee members along with evaluation and
implementation research to develop questions to measure the various aspects of each construct.
Ten questions from the Fisicaro study were included within the survey and other questions were
developed to align with Michigan's mandates of evaluation and the constructs within the study.
Table 3.1 presents an overview of the constructs and questions related to evaluation literature
and research. Table 3.1 presents the original three constructs and questions that were included
within the principal survey prior to factor analysis.

Table 3.1. Constructs and Questions

Construct	Questions	Support from Literature
Clear and Rigorous Expectations	I have a clear picture of the criteria that is used to evaluate teachers.	New Teacher Project, 2010 The Education Trust-Midwest, 2011 Fisicaro, 2010
	I have a clear understanding of the definition of effective teaching.	Danielson & McGreal 2000
	I have a clear understanding of the rubric language.	The Danielson Group, 2011
	I believe principals should provide teachers with advance notice of evaluative classroom visits.	Glickman, 2001 Dudney, 2002 Fisicaro, 2010
	I believe principals should conduct pre-conferences with staff members prior to evaluative visits.	Danielson & McGreal 2000 Fisicaro, 2010

	I believe principals should evaluate teachers at least annually.	New Teacher Project 2010
Involve Multiple Measures	I believe teacher evaluation should include multiple sources of data.	Little, Goe, & Bell, 2009 Weisberg, Sexton, Mulhern & Keeling, 2009 Fisicaro, 2010
	I believe individual student growth should be measured as a part of teacher evaluation.	Little, Goe, & Bell, 2009 Fisicaro, 2010
	I believe teacher evaluation should rely on data collected throughout the school year rather than one moment in time.	Danielson & McGreal 2000 MET, 2012
	I believe information gained through regular contact with staff should be used in evaluations.	Schmoker, 2003 MET, 2012 Fisicaro, 2010
	I believe teacher evaluation should incorporate feedback from students and parents.	Little, Goe, & Bell 2009
	I believe relevant accomplishments and contributions to the school should be considered in teacher evaluation.	Little, Goe, & Bell, 2009 RAND, 2011
	I believe student assessment should be utilized as a measure of teacher effectiveness.	Tucker & Strong, 2005 Fisicaro, 2010
	I believe principals should not evaluate teachers on the basis of one or two classroom visits.	Weisberg, Sexton, Mulhern & Keeling, 2009 Fisicaro, 2010

80

	I believe principals should not ask teachers to provide evidence of student learning.	Little, Goe, & Bell, 2009 Fisicaro, 2010
	I believe principals should not include a teacher's attendance as part of the evaluation.	
Support Ongoing Educator Learning And Development	I believe the principal's primary role of instructional supervision should be to improve instruction.	Davis, Ellet, & Annunziata, 2004
	I believe the role of a principal in the teacher evaluation process should be more closely aligned with that of a coach.	Danielson, 2000 Fisicaro, 2010
	I believe principals should conduct follow up meetings to provide individual feedback following a classroom observation.	Danielson & McGreal 2000
	I believe I am able to facilitate development of professional growth plans based on information from teacher evaluations.	Little, Goe, & Bell, 2009
	I believe our new teacher evaluation is an effective process for evaluating teacher performance.	Little, Goe, & Bell, 2009
	I believe our district evaluation tool will assist ineffective teachers to make significant changes in their instructional practice.	Boyd, 1989 Little, Goe, & Bell, 2009
	I believe our new teacher evaluation process is an effective tool to help teachers improve their performance.	Danielson & McGreal 2000
	I believe a teacher's special training beyond what is required by the district and continuing education should be considered in the evaluation process.	Little, Goe, & Bell, 2009
	I believe a teacher's professional growth plan is an effective tool to improve his or her professional practice.	Boyd, 1989

Pilot Study

Prior to administering the survey the instrument was field-tested with a focus group of five administrators to ensure validity. The survey questions were entered in Qualtrics and administered via email. The pilot study allowed for final opportunity to refine survey questions to ensure clarity of questions, prior to administering the survey to the entire sample of principals and vice principals within the state of Michigan. Creswell (2009) recommended a pilot or field test to "establish the content validity of an instrument and to improve questions, format, and scale" (p. 150). Five principals at the primary and secondary level were included in the pilot study: two elementary principals, one middle school principal, and two high school principals: one vice-principal and one principal. I administered the survey to the principals and then met with each principal to check for question clarity, overall survey feasibility, and any recommendations based on their experience.

Following the pilot study I analyzed the responses and made revisions based on the principals' feedback and dissertation committee member feedback. Two of the principals recommended clarification in regard to the definition of formal and informal teacher evaluation training. This was addressed in the survey by including specific examples of both formal and informal training. All principals highlighted the necessity to increase the scale in the question regarding how much time per year was spent on teacher evaluation to include a choice higher than 100+ hours. Along with these recommendations, the dissertation committee recommended revisions to the survey to include a sliding scale based on a number system, rather than the five options from strongly disagree to strongly agree, as well as a fill in the blank option for the number of hours spent on teacher evaluation to provide a more accurate reporting of time spent based on individual experience.

The pilot study and committee members' recommendations allowed for final survey revisions to establish the content validity of the survey instrument and to improve questions and question response scales prior to administering the survey to the sample of principals and vice principals within the state of Michigan.

Survey Administration

Following the approval of Institutional Review Board, the survey questions were entered into the web-based software program of Qualtrics and administered via e-mail to 1200 public school principals and vice principals within the state of Michigan. Michigan public schools principals' e-mails were found using the Educational Entity Master list found on the Center for Educational Performance website. The survey included additional questions to collect demographic information that was utilized in the descriptive statistics and data analysis. After the final e-mail reminder the survey was closed and the data from Qualtrics was imported to SPSS to conduct the statistical analysis.

Once the survey data were collected and organized, the researcher conducted a factor analysis of the constructs and questions to determine if the questions were valid within the constructs in the survey. The researcher conducted factor analysis on the 25 questions included within the survey and the original three constructs. The factor analysis identified seven constructs within the survey, rather than the original three.

The factor analysis was conducted to determine the extent to which measures overlap or exist among the data in order to create components or constructs within the data set. The initial extraction of variables conducted using the extraction of principal components indicated eight separate components for the independent variables within the research study. Upon further analysis using the principal component analysis and the rotation method promax with kaiser

normalization the results indicated seven components within the data set. The following

constructs were identified within the factor analysis: clear and rigorous expectations (Rig)

multiple measures (MM), feedback (feedback), instructional role of principal (InstruRole), other

evidence (Evidence), professional development (ProfDev), and informal multiple measures

(InfMM).

Factor analysis was conducted to identify factors that explain common variance among

variables and to reduce data by grouping variables that measure a common construct (Mertler &

Vannatta, 2010). The following questions were found to not be valid measures within the

constructs based on the analysis: I believe principals should NOT evaluate teachers on the basis

of one or two classroom visits. I believe principals should evaluate teachers at least annually. I

believe a teacher's professional growth plan is an effective tool to improve his or her

professional practice. Questions were grouped based on validity scores and these questions did

not indicate a high validity score within any of the seven constructs. Based on the analysis these

three questions were found to not be valid measures of any of the seven constructs identified

within the analysis, so these questions were not included within the final analysis of the study.

Table 3.2 reports the constructs based on the outcome of the analysis.

Table 3.2. Factor Analysis

	Rig	MM	Feedback	InstruRole	Evidence	ProfDev	InfMM
I believe our new teacher evaluation process is an effective tool to help teachers improve their performance. (Q86)	.809						
I believe our new teacher evaluation is an effective process for evaluating teacher performance. (Q85)	.780						
I believe our district evaluation tool will assist ineffective teachers to make significant changes. (Q84)	.762						
I have a clear picture of the criteria that is used to evaluate teachers. (Q61)	.696						
I have a clear understanding of the rubric language. (Q63)	.678						
I believe I am able to facilitate development of professional growth plans based on information from evaluation. (Q82)	.606						
I have a clear understanding of the definition of effective teaching. (Q62)	.534						
I believe a teacher's professional growth plan is an effective tool to improve his or her professional practice. (Q83)	.501						
I believe individual student growth should be measured as part of teacher evaluations. (Q70)		.778					
I believe student assessments should be utilized as a measure of teacher effectiveness. (Q75)		.685					

I believe principals should conduct follow up meetings to provide individual feedback following an evaluation. (Q81)	.677		
I believe principals should conduct pre conferences with staff members prior to evaluative visits. (Q65)	.572		
I believe teacher evaluations should include multiple measures. (Q69)	.560		
I believe principals should provide teachers with advance notice of evaluative visits. (Q64)	.508		
I believe teacher evaluation should incorporate feedback from parents.(Q73_2)		.861	
I believe evaluation should incorporate feedback from students. (Q73_1)		.850	
I believe the principal's primary role of instructional supervision should be to improve instruction. (Q79)		.718	
I believe the role of the principal in the teacher evaluation process should be closely aligned with a coach. (Q80)		.715	
I believe principals should not ask teachers to provide evidence of student learning. (Q77)			.655
I believe principals should NOT include a teacher's attendance as part of the evaluation. (Q78)			.634
I believe relevant accomplishments and contributions should be considered in teacher evaluation. (Q74)			.726

I believe a teacher's special training beyond what is required by the district and continuing education. (Q87)	.636
I believe information gained through regular contact with staff should be used in evaluations. (Q72)	.645
I believe teacher evaluations should rely on data collected throughout the school year rather (Q71)	.587

The factor analysis identified seven constructs within the study, including: clear and rigorous expectations, multiple measures, informal multiple measures, feedback, instructional role of the principal, evidence, and professional development. Following the factor analysis a reliability analysis was performed to determine if the questions were accurate measures of the constructs. The reliability analysis indicated that the questions consistently reflected the constructs being measured in five of the seven constructs: clear and rigorous expectations, multiple measures, feedback, instructional role of the principal, and professional development. Clear and rigorous expectations and feedback had outputs higher than .8 and multiple measures had an output of .7. Instructional role of the principal and professional development indicated a weaker Cronbach's alpha but were still included within the study. The Cronbach's alpha indicated overall reliability of the questions in the survey and the constructs, with the exclusion of the constructs of evidence and informal multiple measures due to their weak score.

Variables

This section outlines the variables within the research questions of the study. In question number one the independent variables include evaluation training, position (principal or vice-principal), type of school, years of service, prior teaching experience, and gender. Years of

87

service, prior teaching experience, and evaluation training formal and informal are continuous variables based on a range. The position of the principal will be a categorical variable including three categories to determine the principal's administrative level: elementary, middle, and high school. Gender and position (principal, vice-principal) were also categorical variables. The dependent or outcome variable were perceptions of the principals reflected in the completed survey.

In question number two the independent variables include perceptions, evaluation training, position (principal or vice-principal), type of school, years of service, prior teaching experience, and gender. Years of service, prior teaching experience, evaluation training, and perceptions were continuous variables within the study. Position of the principal, type of school and gender were categorical variables. The dependent variable was implementation of the formal teacher observation and informal teacher observations reflected in the survey.

Analytical Method

The data analysis aimed to measure principals' perceptions of the constructs based on the new teacher evaluation mandates in Michigan, and to determine if the defined variables impacted local implementation. The data analysis sought to reveal whether or not principals' perceptions, position, years of service, gender, prior teaching experience, and level of evaluation training impacted the implementation of formal teacher observations and informal teacher observations. The analytical method is outlined based on the two research questions within the study.

A multiple regression test was used for analysis for question one in the research study. The first question included one outcome variable, perceptions of principals, which is continuous, and two or more predictor variables, which include years of service, level of school, teaching experience, evaluation training, position, and gender, both categorical and continuous predictor

variables, indicating a multiple regression test for analytical method of study. A multiple

regression test according to Field (2009) is an extension of a simple regression in which outcome

is predicted by linear combination of two or more predictor variables.

1. To what extent does a principal's years of service, position (principal or vice-principal), type of school, prior teaching experience, evaluation training, and gender influence their perceptions in regard to the new teacher evaluation mandates in the state of Michigan?

Independent Variable- Years of service, position (principal or vice-principal), type of school, prior teaching experience, evaluation training, and gender
Dependent Variable- Perceptions

Multiple regression was also used for the second question, based on the outcome variable

of implementation as a continuous variable, two or more predictor variables including years of

service, level of school, teaching experience, perceptions, evaluation training, position, and

gender, both categorical and continuous variables.

2. To what extent does a principal's years of service, position (principal or vice-principal), type of school, perceptions, prior teaching experience, evaluation training, and gender impact the implementation of formal teacher observation and informal teacher observation in response to the new teacher evaluation mandates in the state of Michigan?

Independent Variable- Years of service, position (principal or vice-principal), type of school, perceptions, prior teaching experience, evaluation training, and gender
Dependent Variable: Implementation of Teacher Evaluation

Methodology Limitations

This quantitative study was delimitated to public school principals in the state of

Michigan, limiting the population to principals and vice-principals involved in the teacher

evaluation process. Due to the nature of the evolving mandates and the continuous changes in

regard to the teacher evaluation process in Michigan, the survey was administered within the

2012-2013 school year, prior to the 2013 recommendation for a statewide educator evaluation

system. The survey was administered at the end of the 2013 school year, which may have

limited the response rate due to the busyness of principals at the end of the year. The survey questions were entered into Qualtrics, administered via e-mail, which may have also limited the response rate.

Summary

The quantitative study aimed to measure principals' perceptions of the new teacher evaluation mandates within Michigan. The study sought to provide insight on how principals perceived the new teacher evaluation mandates to determine if the defined variables impacted local implementation in an attempt to assist stakeholders in teacher evaluation reform.

CHAPTER IV

PRESENTATION OF RESULTS

Introduction

Chapter four presents the results from the analysis of the study described in chapter three. This chapter includes the descriptive statistics for the independent continuous variables, the frequencies of the independent categorical variables, and the multiple regression analysis for both research questions. A multiple regression was used for both research questions to determine if the independent variables predicted the dependent variables within the study.

> 1. To what extent does a principal's years of service, position (principal or vice-principal), type of school, prior teaching experience, evaluation training, and gender influence their perceptions in regard to the new teacher evaluation mandates in the state of Michigan?

> 2. To what extent does a principal's years of service, position (principal or vice-principal), type of school, perceptions, prior teaching experience, evaluation training, and gender impact the implementation of formal teacher observation and informal teacher observation in response to the new teacher evaluation mandates in the state of Michigan?

Results

Descriptive Statistics

Descriptive statistics were analyzed and all variables were checked for normality. Some variables demonstrated a skewness and kurtosis that required an adjustment for normality. The variables corrected included evaluation training formal hours and evaluation training informal hours, both variables indicated a positive skew. A positive skew indicates a large number of participants reported on the low end of the scale for hours of formal and informal training. The logarithms of evaluation training both formal and informal were created to adjust for normality.

The analysis indicated that the log bas ten for some values was less than or equal to zero on the indicated command and the result was set to a system-missing value.

Outliers were identified within the data and corrected prior to further analysis. Perceptions, years of service, and implementation contained outliers beyond three standard deviations of the mean and these values were adjusted prior to further analysis. Administrator perceptions was found to have an outlier of the minimum, so the data value of 77 was changed to within three standards deviations of the minimum value of 81. One value for an administrator's years of service was found to be an outlier of the maximum with 40 years of service reported, so the years were adjusted to 32 to bring the data point within the maximum range. Implementation was also found to have two outliers below the minimum value, so 14 was changed to 30 and 24 was changed to 30. All outliers were brought within three standard deviations of the mean. Linearity was checked and no corrections were necessary.

Following the correction, descriptive statistics was run again to report the new outcomes for mean, standard deviation, skewness, and kurtosis to represent the correction as outlined in table 4.1.

Table 4.1. Descriptive Statistics of Construct 1, Construct 2, Construct 3, Construct 4, Construct 6, Years of Service, Prior Teaching Experience, Evaluation Training Formal, Evaluation Training Informal, Perceptions, and Implementation

	⌣	N	Mean	Std. Deviation	Skewness	Kurtosis
Construct 1: Clear and Rigorous Expectations	.83	180	55.011	12.043	-0.469	-0.251
I believe our new teacher evaluation process is an effective tool to help teachers improve their performance.		180	6.326	2.465	-.679	-.171
I believe our new teacher evaluation is an effective process for evaluating teacher performance.		180	6.149	2.373	-.661	.211
I believe our district evaluation tool will assist ineffective teachers to make significant changes.		180	6.034	2.261	-.514	-.412
I have a clear picture of the criteria that is used to evaluate teachers.		180	7.283	2.064	-.722	-.037
I have a clear understanding of the rubric language.		180	7.257	2.310	-1.047	.765
I believe I am able to facilitate development of professional growth plans based on information from evaluation.		180	7.067	1.906	-.746	.876
I have a clear understanding of the definition of effective teaching.		180	8.217	1.642	-1.379	2.802
I believe a teacher's professional growth plan is an effective tool to improve his or her professional practice.		180	7.226	2.060	-.679	-.166
Construct 2: Multiple Measures	.67	180	41.83	8.483	-0.416	0.78
I believe individual student growth should be measured as part of teacher evaluations.		180	7.361	2.310	-.932	.656

		N	Mean	SD	Skew	Kurt
I believe student assessments should be utilized as a measure of teacher effectiveness.		180	6.978	2.399	-.931	.527
I believe principals should conduct follow up meetings to provide individual feedback following an evaluation.		180	8.715	1.700	-1.433	1.902
I believe principals should conduct pre conferences with staff members prior to evaluative visits.		180	6.486	3.064	-.485	-.965
I believe teacher evaluations should include multiple measures.		180	9.068	1.308	-1.818	4.166
I believe principals should provide teachers with advance notice of evaluative visits.		180	3.659	2.803	.470	-.596
Construct 3: Feedback	.84	180	10.828	5.332	-0.212	-0.775
I believe teacher evaluation should incorporate feedback from parents.		180	5.11	2.884	-.150	-.876
I believe evaluation should incorporate feedback from students.		180	5.869	2.897	-.337	-.839
Construct 4: Instructional Role of Principal	.48	180	16.829	2.832	-1.109	1.245
I believe the principal's primary role of instructional supervision should be to improve instruction.		180	8.856	1.304	-1.366	2.077
I believe the role of the principal in the teacher evaluation process should be closely aligned with a coach.		180	8.017	2.059	-1.346	1.962
Construct 6: Professional Development	.52	180	15.178	3.631	-1.00	1.533
I believe relevant accomplishments and contributions should be considered in teacher evaluation.		180	8.196	1.8022	-1.467	3.328

I believe a teacher's special training beyond what is required by the district and continuing education.	180	7.107	2.418	-.825	.052
Years of Service	176	10.89	7.402	0.900	0.390
Teaching Experience	177	3.45	1.201	0.383	-0.535
Evaluation Training Formal	180	16.45	23.224	4.356	27.288
Evaluation Training Formal Logarithm	154	1.0749	0.435	-0.135	0.52
Evaluation Training Informal	175	15.82	19.328	2.594	8.328
Evaluation Training Informal Logarithm	173	0.9929	0.447	0.034	-0.359
Perceptions	180	159.844	21.287	-0.432	0.034
Implementation	180	69.274	12.866	-0.619	0.339
Pre-Observation Conference	180	7.812	2.971	-1.408	.895
Classroom Walk Through	180	8.819	2.155	-2.549	6.526
Informal Conversations	180	8.066	2.431	-1.599	2.115
Evaluation Rubric	180	9.448	1.301	-3.763	18.994
Observation Checklist	180	8.000	2.716	-1.638	1.962
Post Observation Conference	180	8.907	2.037	-2.456	5.961
Teacher Professional Growth Plan	180	8.105	2.376	-1.465	1.735
Teacher Portfolio	180	5.054	3.341	-.017	-1.123
End of Year Summative Narrative	180	7.854	3.109	-1.455	.868

Table 4.2 identifies the frequencies of the categorical independent variables within the research study: position, gender, and level of school. The frequency table identifies the difference in principal and vice-principal response rate. A greater amount of principals (136)

responded to the survey than vice principals (44). The gender ratio of participants was more even with 97 males and 82 females participating in the study. Elementary, middle, and high school principals were fairly equally represented in the study as well.

Table 4.2. Frequencies of Position, Gender, and Level of School

		Frequency	Percent
Position	Principal	136	75.6
	Vice Principal	44	24.4
Gender	Male	98	54.4
	Female	82	45.6
Level of School	Elementary	60	33.3
	Middle School	56	31.1
	High School	64	35.6

Multiple Regression Question One

For question number one a multiple regression was conducted to determine if the independent variables (years of service, level of school, prior teaching experience, formal and informal evaluation training, gender, and position) influence a principal's perceptions of the new teacher evaluation mandates. The survey assessed principals' perceptions of the expectations of the new teacher evaluation process and their role in the course of evaluation.

Principals were asked questions in regard to the new system to determine if they perceived the new process as a proficient means to evaluate teacher effectiveness and encourage student growth, as well as their perceptions of the process and their understanding of the rubric language, criteria, and expectations under the new law. For question number one, the survey

aimed to determine if principals' perceptions were impacted based on the variables identified within the study.

The model results show that the independent variables within the study do not predict the dependent variable: principal perceptions. A non-significant F statistic demonstrated that the model is not an accurate prediction of the outcome variable within the study. The adjusted R squared (.013) indicated that only 1.3 percent of the perceptions of principals is explained by a principal's years of service, level of school, prior teaching experience, formal and informal evaluation training, gender, and position.

The survey measured principals' perception of the new evaluation system and the components included within the new evaluation system. Principals agreed and strongly agreed that the new evaluation should include principals in an instructional role, clear and rigorous expectations, multiple measures, feedback, and ongoing professional development as a means to a more comprehensive teacher evaluation, but did not strongly agree that the new evaluation tool was the most effective tool in evaluating teacher performance or strongly agree that the new evaluation is an effective tool to help teachers improve their performance.

Principals consistently agreed that the evaluation process should have clear and rigorous expectations. The data indicated that principals believe they have a clear understanding of the criteria used to evaluate teachers, definition of effective teaching, and rubric language. Principals strongly agreed that evaluations should include multiple measures, but only agreed rather than strongly agreed that individual student growth and student assessments should be utilized as measures in the evaluation process. Principals also agreed that relevant accomplishments, contributions, and a teacher's special training beyond the district's requirements should be considered in the evaluation process.

97

Principals also consistently agreed that their role in the evaluation process should be more of a coach aimed to improve instruction. This is consistent with their rating in regard to providing feedback following the evaluation. The survey indicated that principals believe they should conduct follow up meetings to provide individual feedback following an evaluation. Principals did not strongly agree or agree that teacher evaluations should incorporate feedback from parents and students in the evaluation process.

Table 4.3 indicates the outcome of the multiple regression analysis for question one.

Table 4.3. Multiple Regression Question One Years of Administrative Service, Level of School, Prior Teaching Experience, Evaluation Training Formal, Evaluation Training Informal, Gender, and Position

Variable	Beta	SE
Years of Administrative Service	-0.244	0.230
Level of School	-0.1446	2.230
Prior Teaching Experience	0.175	1.416
Evaluation Training Formal	11.198	4.492
Evaluation Training Informal	-0.527	4.416
Gender	0.75	3.610
Position	1.500	4.242

Adj R^2 = .013, F = .271

***$p < .01$, **$p < .05$

After running the multiple regression the error term was examined for auto correlation and homoscedasticity by examining a scatter plot of the predicted variable against the residual variable and no relationship was found.

Multiple Regression Question Two

For question number two a multiple regression was conducted to determine if the independent variables (years of service, level of school, prior teaching experience, formal and informal evaluation training, perceptions, gender, and position) influenced the implementation of formal and informal teacher observation in response to the new teacher evaluation mandates. The survey asked principals to indicate how likely they were to use formal and informal observation instruments and procedures to evaluate teachers under the new law.

The survey incorporated formal observation instruments and procedures including pre and post conferences, evaluation rubrics, observation checklists, teacher professional growth plans, teacher portfolios and summative narratives, as well as informal methods including classroom walkthroughs and conversations with teachers. Principals indicated that they were likely to incorporate both formal and informal evaluation procedures in the process of evaluation to include pre and post conferences, classroom walkthroughs, informal conversations, rubrics, checklists, teacher professional growth plans, and summative narratives. Principals indicated that they were not as likely to implement teacher portfolios in the evaluation process.

The model results show that the independent variables within the study do not have a significant relationship in predicting the dependent variable of teacher evaluation implementation. A non-significant F statistic indicated the model is not an accurate prediction of the outcome variable within the study. Adjusted R squared (.046) indicated that only about five percent of implementation is influenced by the variables within the study: years of service, level of school, prior teaching experience, perceptions, formal and informal evaluation training, gender, and position.

99

Table 4.4 presents the outcome of the multiple regression analysis for question two in the research study to include the beta and standard error value.

Table 4.4.Multiple Regression Question Two Years of Administrative Service, Level of School, Prior Teaching Experience, Evaluation Training Formal, Evaluation Training Informal, Gender, Position, and Perceptions

Variable	Beta	SE
Years of Administrative Service	-0.171	0.145
Level of School	1.694	1.385
Prior Teaching Experience	1.591	0.884
Evaluation Training Formal	2.959	2.844
Evaluation Training Informal	-0.713	2.740
Gender	-0.683	2.260
Position	-0.136	2.645
Perceptions	0.114	0.054

Adj R² = .046, F = 0.74

***p < .01, **p <.05

Following the multiple regression analysis the error term was examined for auto correlation and homoscedasticity by examining a scatter plot of the predicted variable against the residual variable and no relationship was found.

Summary

Chapter four included the descriptive statistics for the independent continuous variables within the study, the frequencies of the independent categorical variables, and the multiple regression outcomes for both research questions.

The p value in both questions was greater than 0.05 indicating that the independent variables are not predictors of either outcome variable: principal's perceptions or formal and informal observation implementation. The multiple regression analysis reported an F statistic that indicated the model is not an accurate prediction of the outcome within the study.

The analysis of the study is reported in chapter four with little or no significance found between the independent variables and the outcome variables in either research question. The researcher will further analyze the findings of the study in chapter five, as well as implications for practice, policy, and future research.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

Teacher evaluation is at the forefront of educational reform due to increased accountability measures at the national and state level. The most recent evaluation reform measures are impacting state and local policy across the United States. The State of Michigan is currently involved in the process of evaluation reform with final recommendations for a statewide evaluation system recently presented by the Michigan Council for Educator Effectiveness (2013) for consideration at the legislative level.

Michigan is currently in the preliminary stages of implementation of a statewide evaluation system and as districts transition toward a more rigorous evaluation system, principals will be required to evaluate all teachers every year. The role of the administrator in the evaluation process is even more critical as accountability measures heighten with high-stakes decisions based on evaluation results. Teacher evaluation is a hot topic across the country with reform measures occurring at a rapid pace. With evaluation results tied to high-stakes decisions it is critical that policy, practice, and research are aligned to ensure that evaluation reform is focused on increasing teaching and learning and implemented with integrity and rigor.

This quantitative study surveyed principals to assess their experiences in response to the new teacher evaluation mandates to determine if the defined variables influenced principals' perceptions and impacted local implementation in public schools across Michigan. Chapter five includes the limitations, interpretations of findings, implications for policy, practice, and future research.

Limitations

This quantitative study was delimited to public school principals in the state of Michigan, limiting the population to principals and vice-principals involved in the teacher evaluation process. Although the findings of this study are specific to the public schools within Michigan, teacher evaluation is a common measure of reform across school districts in the United States. Reforming teacher evaluation is a complicated process and many states have implemented new systems for evaluation, making it necessary for states to continue research and to conduct studies regarding evaluation and implementation to broaden the findings.

Due to the nature of the evolving mandates and the continuous changes in regard to the teacher evaluation process in Michigan, the survey needed to be administered within the 2012-2013 school year, prior to the 2013 final recommendation for a statewide educator evaluation system. The survey was administered in May 2013 at the end of the school year. The timing for the survey was ideal in the sense that most administrators were fully engaged in the process of evaluation. The new teacher evaluation mandates in Michigan require principals to evaluate all teachers every year and to report yearly effectiveness ratings to the State of Michigan registry at the end of each school year.

The survey aimed to assess principal's perceptions of the mandates at a time in which they were fully engaged in the process of evaluation and implementation. Although this was considered an ideal time for administering the survey to capture principal's perceptions, the timing could also be considered problematic in the sense that administrators were too busy to respond to the survey due to involvement in the end of the year administrative duties within their schools. The survey was administered via e-mail through Qualtrics, which may have also limited the response rate.

The following section includes the interpretation of the findings of both research

questions and the implications for policy, practice, and future research.

Interpretations of Findings

The following section outlines each research question in the study and the interpretations

of findings.

1. To what extent does a principal's years of service, position (principal or vice-principal), type
of school, prior teaching experience, evaluation training, and gender influence their perceptions
in regard to the new teacher evaluation mandates in the state of Michigan?

Principals' perceptions were assessed through their responses to nineteen questions

based on the new teacher evaluation mandates in Michigan. Demographics were also collected

for analysis to include a principal's years of service, type of school, number of hours for both

formal and informal evaluation training, number of years as a classroom teacher, position and

gender to determine if these variables influenced principal's perceptions of the new teacher

evaluation mandates.

Principals were asked questions based on the new evaluation mandates and the design

principles outlined within the new mandates for a more comprehensive teacher evaluation

system: clear and rigorous expectations, multiple measures, and ongoing learning and

development. The clear and rigorous expectations of the system intend to provide principals

with an effective tool and process to evaluate teacher performance and to encourage instructional

growth of teachers. Principals were also surveyed in regard to an evaluation tool that considers

multiple measures within the process of evaluation that includes student growth, feedback from

parents and students, and professional development of teachers. The results of the survey

indicated that principals' perceptions in regard to the new teacher evaluation mandates were not

influenced by a principal's background, experiences, or access to evaluation training.

Different groups of principals did not indicate differences in perceptions of the new mandates, meaning that perceptions of the mandates were generally consistent within the study. Although principals rated several questions within the constructs as valid measures within the new evaluation process, the overall perception of the new evaluation tool was rated lower. Principals agreed and strongly agreed that the new evaluation should include principals in an instructional role, clear and rigorous expectations, multiple measures, feedback, and ongoing professional development as a means to a more comprehensive teacher evaluation, but did not strongly agree that the new evaluation tool was the most effective tool in evaluating teacher performance or strongly agree that the new evaluation is an effective tool to help teachers improve their performance.

This outcome could be due to various circumstances: principals' perceptions of teacher evaluation and the necessary elements in a more comprehensive process are aligned with modern evaluation research, principals are simply abiding by the new mandates based on the law, or the implementation process is too preliminary to accurately gauge principals' perceptions and implementation of the new mandates.

Teacher evaluation reform has created a demand for a more comprehensive approach to evaluation with principals as the primary implementers. With reform measures happening so rapidly at the national and state level, principals must remain aware of current theory, policy, and practice in the educational realm and the major paradigms impacting reform. Effective principals are closely involved in the district evaluation process and are aware of the valuable measures in evaluation through continued practice and efforts to stay current with modern evaluation research (Goe, Bell, & Little, 2008; Little, Goe & Bell, 2009; MET, 2013; The New Teacher Project, 2010).

Evaluation reform has been evolving in Michigan for several years, with principals at the forefront of the change. The new evaluation system in Michigan was designed based on evaluation research and applied practice. Modern evaluation research outlined key components necessary for inclusion in the development of a more comprehensive teacher evaluation system aimed to drive instructional improvement to include clear and rigorous expectations, multiple measures, and measures to support educator learning and development (Daley & Kim, 2010; Goe, Bell, & Little, 2008; Little, Goe & Bell, 2009; MET, 2013; The New Teacher Project, 2010; Weisberg, Sexton, Mulhern, & Keeling, 2009). Principals in Michigan perceived several of the components in the new teacher evaluation as a means to an improved evaluation system, which could be contributed to their understanding of modern evaluation practice and research.

This study indicated that principals support several of the components in the new teacher evaluation process as means to an improved evaluation system, but do not strongly agree that the new evaluation process and tool are the most effective measure of teacher effectiveness. The new evaluation process is only in the preliminary stages of implementation, with final recommendations not yet implemented in local districts. Principals' perceptions of the new tool may change once the new tool is fully implemented across local districts in Michigan.

The MCEE (2013) statewide evaluation recommendations aimed to develop an improved system designed to support efforts toward improvement rather than compliance. Fisicaro argued (2010) that a principal's attitudes toward the necessary components of evaluation will determine the levels to which principals commit to the application of the research. This study did not provide evidence as to whether or not principals were likely to implement these practices because they agreed they were important in the evaluation process or if they simply implemented the practices because of the new mandates.

106

The findings of the study indicated that perceptions of the new system were not impacted by a principal's years of service, type of school, position, prior teaching experience, evaluation training, and gender, indicating that principals shared similar views pertaining to the new evaluation system, the expectations of the system, and the criteria for evaluating teachers. The findings of the study indicated that the different attributes of principals did not impact their perceptions of the new evaluation mandates. These findings are noteworthy because the MCEE included these as recommended measures for evaluating teacher effectiveness based on the pilot study, evaluation research, and consultation. The findings of this study are important to the ongoing efforts to encourage implementation of the mandates with consistency across Michigan, regardless of the school district or evaluator. The following section will outline the second research question within the study.

2. To what extent does a principal's years of service, position (principal or vice-principal), type of school, perceptions, prior teaching experience, evaluation training, and gender impact the implementation of formal teacher observation and informal teacher observation in response to the new teacher evaluation mandates in the state of Michigan?

The second question assessed the process of implementation and the likeliness of implementation of formal and informal observation practices based on the new mandates. This question sought to determine if the defined variables impacted the implementation of formal and informal teacher observation practices in response to the new mandates in Michigan. Demographics were also used for analysis in the second question to include a principal's years of service, type of school, number of hours for both formal and informal evaluation training, number of years as a classroom teacher, position, and gender. The final analysis indicated that the model was not an accurate predictor of the outcome variable.

Principals were surveyed regarding implementation of both formal and informal observation methods in the evaluation process. Principals reported consistently that they were likely to implement both formal and informal observation practices including observation conferences pre and post, classroom walk through, informal conversations, evaluation rubrics, observation checklists, teacher professional growth plans, and summative narratives in the evaluation process. This is also noteworthy in the evaluation process because the MCEE (2013) included these as recommended measures for evaluating teacher effectiveness. Principals in Michigan are required to implement the new teacher evaluation process and prior to implementation, principals must first understand and value the elements and dimensions that modern evaluation systems require.

The study indicated that principals were likely to implement the formal and informal measures in the process of evaluation. This could also be due to similar circumstances as question one within the study: principals' are aware of the new mandates and requirements in the evaluation process and are simply abiding by the mandates or principals believe the recommended formal and informal measures are accurate measures in the evaluation process. Again this study did not provide evidence to determine whether principals are responding based on their professional beliefs or if their responses are reflective of the new law.

This study was conducted prior to full implementation of the statewide educator evaluation system, and I believe it is important to consider these variables in future studies to again assess significance once the evaluation system has been fully implemented within local districts. The data indicated that both formal and informal training was skewed due to a large number of participants that reported on the low end of the scale. Due to the preliminary stage of implementation, evaluation training both formal and informal can not be accurately accessed

based on the survey data, due to the majority of participants receiving little or no training in regard to the new evaluation tool. With teacher evaluation as a means for evaluating teachers with decisions related to high-stakes measures, it is imperative that continued efforts and studies are conducted in regard to the new evaluation system, a principal's role in the process, and the integrity of implementation.

The following section will discuss the study findings and implications for policy, practice, and future research.

Implications

This section will discuss how the results of this study are important in the preliminary stages of evaluation reform and the continued efforts to ensure successful policy implementation in local schools.

The findings of the study indicated that years of service, type of school, position, prior teaching experience, evaluation training, and gender did not have a significant impact on a principal's perceptions of the new evaluation mandates or the implementation of the mandates. Principals indicated consistent ratings in regard to their perceptions and implementation of the new mandates. The new evaluation mandates outline the requirements for the new statewide evaluation system for implementation in the local districts. As districts proceed with implementation, it is critical that the system is continually monitored to ensure reliable implementation.

The findings of this study are significant to the ongoing efforts to ensure implementation of the mandates with consistency across Michigan, regardless of the school district or evaluator. This study did not take into consideration other factors that could impact implementation such as the size of a school, the number of teachers a principal is responsible for evaluating, time spent

on evaluation, access to data-systems, or the leadership style of the evaluator. Future studies should be conducted at the national and state level to identify factors that influence policy implementation in local districts based on continued educational research and applied practice.

Implications for Policy

      With the continual push for increased accountability and educational reform in the United States and in individual states, educational policy remains at the forefront of change with national and state policy being created to transform educational practice. Educational reform has proved to be controversial at the federal and state level with a great deal of educational policy influenced by various stakeholders with diverse interests in the educational realm in our country. Educational reform has sparked the interest of many organizations including those in the public and private sector. Reform measures are happening at a rapid pace in this country and national and state policies are being created in response to the various reform initiatives.

      The MCEE (2013) delivered their final recommendations to the State of Michigan workgroup in the fall of 2013 following 18 months of research and collaboration in the development of recommendations for a statewide evaluation system. The workgroup comprised of bi-partisan legislators is currently reviewing the recommendations to inform policy at the state level. The new teacher evaluation system is in the preliminary stages of implementation and it is important that continued efforts are made to monitor, evaluate, and assess the evaluation system in years to come to ensure continued successful implementation and to identify factors that influence implementation.

      The MCEE and State of Michigan are currently in the preliminary stage of implementation of the new statewide evaluation system. The MCEE has recommended an ambitious teacher evaluation system with the true challenges yet to be seen. Once the State

brings forward the MCEE recommendations with legislation to create evaluation policy, the next step is full implementation of the new evaluation system in local districts. The MCEE has recommended a statewide evaluation system and the state will require local districts to put policy into practice. Not only will the state need to continue efforts to evaluate the new system and its outcomes, the state will also need to use teacher evaluation results in meaningful ways to inform policy reform and practice.

Doherty, Jacobs, Lakis, Staresina, and Lasser (2013) stated

> The policy implications of an evaluation system that truly measures teacher effectiveness are profound. If done well, and if decision makers act on the results, the consequences could change much of what is now standard practice in the teaching profession by setting the foundation for better targeted policies for struggling teachers, higher standards for teacher preparation programs and fair but rigorous policies for replacing persistently ineffective teachers. (p.17)

Doherty et al. (2013) argued that the process to reform a state evaluation system is costly and not using the results in meaningful ways is counterproductive and wasteful. The point of improved evaluation should be to improve teacher practice and student learning. States must use the information from teacher evaluations about effectiveness to inform teacher policy and local district practice.

Due to the increase in reform and policy creation, many states are experiencing policy change in regard to teacher evaluation, with many policies not yet in practice (Doherty, et al., 2013). Policy implementation is a complex process in the educational environment and continued efforts are necessary to ensure the policy is meeting the intended outcomes and is implemented with fidelity and rigor.

111

The following section will discuss the implications in the process of implementing policy as new practice within local districts.

Implications for Practice

The new teacher evaluation mandates in Michigan require school districts to adopt and implement a rigorous, transparent, and fair performance evaluation system to evaluate teachers annually to provide constructive feedback to enhance teaching and learning. School districts will be required to adopt the new evaluation systems with responsibility at the local level for implementation of the new practice.

The new teacher evaluation system is only in the preliminary stages with the most recent recommendations not yet fully implemented. Leadership needs to not only focus on the policy, but more importantly on the success of implementation and the factors that impact implementation. The implementation process at the local level is a critical factor in the success of a new teacher evaluation system. Although this study did not find significance among the variables and implementation, it is an area that requires further study following full implementation of the statewide evaluation system. Proper implementation of the new system is crucial to determine if the new evaluation methods are consistent in the effort to enhance teacher performance and student learning.

The Implementation Drivers framework guided this study, even though this framework has not been specifically applied to the implementation process of the new teacher evaluation mandates by the State of Michigan, because they provide a useful way to think about effective policy implementation. The framework identified specific drivers that influence successful policy and program implementation: competency, capacity, and leadership. These drivers are integrated and compensatory aimed to improve a local district's implementation effectiveness.

The survey specifically examined the competency and capacity drivers involved in the implementation process in regard to the new teacher evaluation mandates within Michigan.

Competency drivers are the mechanisms to develop, improve, and sustain an organization's ability in the implementation process through training and coaching. Within the first two years of implementation of the new evaluation law in Michigan, principals received little or no concerted training in conducting evaluations and lack of training was identified as an issue to be recognized when evaluating reported effectiveness ratings (Keesler & Howe, 2012). Training was a variable considered within the study that was found to have no significant impact on implementation. The data for both formal and informal training was found to be skewed within the study due to a large number of participants that reported on the low end of the scale, with many participants that reported zero hours of training in regard to the new evaluation process. Although this study found no relationship among training and implementation, it is necessary to monitor this aspect of the new statewide system to ensure all principals receive adequate training and coaching in order to implement the statewide system with consistency and validity following full implementation.

Providing training for principals and leadership involved in teacher evaluation is crucial in the process of successful implementation (MET, 2013; State of Michigan, 2011). Principals must be trained in the use of the state-selected observation tool and trained to coach and provide feedback in the evaluation process. Formal training provided by the observation tool vendor is important to provide principals with consistent training measures state-wide. Principals should receive training based on the newly recommended tool with opportunity to observe the tool in practice and opportunity to practice using the tool with follow-up training to provide additional support to address questions and concerns regarding implementation. Principals should be

retrained every two to three years to ensure continued reliable implementation. Districts may also find it beneficial to provide principals with opportunity to take part in informal training measures at the district level to afford opportunity for discussion and feedback based on district experiences.

Inter-rater reliability is also an important aspect to be addressed in training to decrease the variability between raters in the evaluation process. The recommended evaluation tool will provide principals with guidelines for evaluation as well as a rubric for use in the evaluation process. Principals need specific training in regard to the use of the rubric and opportunity to practice. Research has found that easily interpreted rubrics along with training are an important element in improving inter-rater agreement and reliability in evaluation (Graham & Westat, 2011).

It is important that educators and evaluators are properly trained with time to acquaint themselves with the tools and "reorient themselves to the changes in their responsibilities that are required by the system" (MCEE, p. 6). Districts not only need training and coaching, but they also need the capacity to implement a more rigorous system.

Capacity is defined as the resources, data systems, and supports in place to make implementation easier for the organization. The MCEE has recommended the State provide evaluation frameworks, data systems, and funding to ensure districts have adequate resources for implementation. This study assessed principals' perceptions of the expectations of the new system to determine if principals believed they had a clear understanding of the criteria of the evaluation process including an understanding of effective teaching and the rubric language incorporated in observation. Districts will be required to implement one of the four observation tools within their districts to evaluate teachers based on the final MCEE recommendations and

114

future efforts will be necessary to monitor and evaluate individual district's capacity for proper implementation.

Principals will be responsible for local implementation of the newly defined teacher evaluation process and their role is crucial in successful implementation. Strong leadership is necessary to provide direction, management, and coordination of the evaluation system. A principal's role in the process of evaluation is imperative to ensure consistency and validity in the outcome of the process (MCEE, 2013). Without effective leadership, the evaluation process may be hindered and the outcomes invalid. "Building administrators are intimately knowledgeable about their district required teacher evaluation systems, they are in unique positions to identify specific policies, practices, and beliefs, which inhibit optimal teacher evaluation effectiveness" (Kersten & Israel, 2005, p.51). Principals are responsible for the management and coordination of the evaluation system to ensure continued successful implementation and to review the progress of the implementation to make adjustments as needed.

With amplified accountability and policy mandates, the educational sector continues to experience increased burden for local policy implementation. The evaluation process will warrant continued efforts by the State and local districts to ensure the evaluation system is implemented with fidelity and rigor, and most importantly to ensure that the system is meeting the intended goals of enhancing instruction, improving student achievement, and supporting professional learning. Educational reform is a complex and expensive process and without proper measures to evaluate effectiveness and the process of implementation, policy will not succeed in practice.

With efforts to develop and implement more rigorous teacher evaluations systems, it is important for continued research efforts in individual states and across the nation as well. Research is a valuable tool to lead states in refining current evaluation policy and practice. The following section will highlight the importance of continued research efforts at the national, state, and local level.

Recommendations for Future Research

Increased accountability in the educational sector has created a sense of urgency for educational policy reform in the United States. The high stakes nature of teacher evaluation has led to increased accountability and decisions based on evaluation outcomes. This increased accountability has led to research efforts to identify more comprehensive teacher evaluation systems that accurately and fairly measure teacher effectiveness and encourage educator and student growth. Comprehensive evaluation frameworks are crucial in the evaluation process, but increased research efforts are necessary that focus on the factors that impact implementation at the local level and efforts that promote sustainability over time.

Numerous initiatives and research partnerships have resulted in the development of practical guides and guidelines for reference in the process of creating more comprehensive and rigorous teacher evaluation systems to evaluate teacher effectiveness in a more precise manner. Evaluation research has outlined key elements necessary for inclusion in the development of comprehensive teacher evaluation systems with more research efforts necessary that focus on sustained policy implementation. Further research is recommended to identify factors that influence implementation and the barriers to successful implementation that exist at the local district level.

The new evaluation system will require continued efforts by the State and local districts to ensure the system is being implemented consistently with fidelity across Michigan. Evaluation research has focused on the essential components necessary for a more comprehensive evaluation framework, with less focus on the factors that impact implementation and ways to evaluate the quality of local implementation. With increased accountability and high-stakes decisions based on evaluation, it is important to focus on the factors that influence implementation to ensure the evaluation system is meeting the intended goals and to inform future reform measures.

> Clearly, creating an educator evaluation system will require developing the relevant policies and processes and creating or adopting relevant instruments to be used. But this system will also require building capacity-both statewide and centrally-to understand, implement, and monitor the system, as well as infrastructure to house data, review materials, produce reports, and conduct program evaluation. (MCEE, 2013, p.30)

The new teacher evaluation system in Michigan is only in the preliminary stage of implementation and the state will be responsible for monitoring compliance with the law and evaluating the quality of local implementation of the law (MCEE, 2013). The MCEE recommended the state be provided adequate resources to monitor compliance and recommended the state conduct a study of the quality of local implementation and the consequences of implementation of a statewide system of evaluation (MCEE, 2013). The MCEE indicated that the study should also include a complete audit of the system after three full years of implementation to examine whether the system has improved teaching and learning in public schools in Michigan.

Continued research efforts are necessary at the national, state, and local level to ensure that new evaluation systems are rigorous, more consistent, and most importantly encourage teacher and student growth. Continued efforts are also necessary to identify the key elements in a more comprehensive teacher evaluation system as well as research efforts that focus on the process of implementation and the factors that encourage or impede implementation.

Conclusions

This study explored principals' perceptions of the new teacher evaluation mandates in Michigan and the impact on implementation. The perceptions of principals were consistent regardless of years of service, type of school, gender, prior teaching experience, evaluation training, perceptions, and position with no significant impact on implementation. Although the study indicated that the variables did not significantly impact perceptions or implementation of the new mandates, it is critical that continued efforts are in place to ensure the new system is suited to the local school district environment, is operating as planned, and most importantly meeting the intended outcomes defined by the new mandates.

With the MCEE recommendations at the legislative level at this time, further research is necessary in order to evaluate the new system after full implementation in districts across the state. Continued efforts are necessary in Michigan and nationwide to identify the most reliable and rigorous methods for inclusiveness in comprehensive teacher evaluation.

APPENDIX

Principal Survey

a. Years of administrative experience

b. Please mark the level of school that you are presently employed

Elementary (K-5)   Middle School (6-8)   High School   (9-12) Other _____

c. How many years of experience do you have as a classroom teacher?

d. How many years experience do you have evaluating teachers?

e. What was the total number of certified teachers within your building in 2011-2012?

f. In the past two years, how many hours of formal evaluation training have you experienced in regard to the new teacher evaluation process?

Lecture Style Training

Practice Using the Tools (Rubric, checklist, etc)

g. In the past two years, how many hours of informal evaluation training have you experienced in regard to the new teacher evaluation process? (Example: Collaboration with administrators)

h.   Gender

Male   Female

SLIDING SCALE

1. Clear and Rigorous Expectations

   1) I have a clear picture of the criteria that is used to evaluate teachers.

   Strongly Disagree   Strongly Agree

   2) I have a clear understanding of the definition of effective teaching.

   Strongly Disagree   Strongly Agree

   3) I have a clear understanding of the rubric language.

   Strongly Disagree   Strongly Agree

   4) I believe principals should provide teachers with advance notice of evaluative classroom visits.

   Strongly Disagree   Strongly Agree

   5) I believe principals should conduct pre-conferences with staff members prior to evaluative visits.

   Strongly Disagree   Strongly Agree

   6) I believe principals should evaluate teachers at least annually.

      Strongly Disagree   Strongly Agree

   7) I am concerned about my inability to manage all that our new teacher evaluation requires.

   Strongly Disagree   Strongly Agree

Open ended question: Please share any specific concerns

2. Involve Multiple Measures

   1) I believe teacher evaluations should include multiple sources of data. (examples)

      Strongly Disagree   Strongly Agree

2) I believe individual student growth should be measured as part of teacher evaluations.

   Strongly Disagree   Strongly Agree

3) I believe teacher evaluation should rely on data collected throughout the school year rather than one moment in time.

   Strongly Disagree   Strongly Agree

4) I believe information gained through regular contact with staff should be used in evaluations.

   Strongly Disagree   Strongly Agree

5) I believe teacher evaluation should incorporate feedback from students and parents.

   Strongly Disagree   Strongly Agree

6) I believe relevant accomplishments and contributions to the school should be considered in teacher evaluation.

   Strongly Disagree   Strongly Agree

7) I believe student assessments should be utilized as a measure of teacher effectiveness.

   Strongly Disagree   Strongly Agree

8) I believe principals should not evaluate teachers on the basis of one or two classroom visits.

   Strongly Disagree   Strongly Agree

9) I believe principals should not ask teachers to provide evidence of student learning.

   Strongly Disagree   Strongly Agree

10) I believe principals should not include a teacher's attendance as part of the evaluation.

   Strongly Disagree   Strongly Agree

3. Support Ongoing Educator Learning and Development

1) I believe the principal's primary role of instructional supervision should be to improve instruction.

   Strongly Disagree   Strongly Agree

2) I believe the role of a principal in the teacher evaluation process should be more closely aligned with that of a coach.

   Strongly Disagree   Strongly Agree

3) I believe principals should conduct follow up meetings to provide individual feedback following a classroom observation.

   Strongly Disagree   Strongly Agree

4) I believe I am able to facilitate development of professional growth plans based on information from teacher evaluations.

   Strongly Disagree   Strongly Agree

5) I believe our new teacher evaluation is an effective process for evaluating teacher performance.

   Strongly Disagree   Strongly Agree

6) I believe our district evaluation tool will assist ineffective teachers to make significant changes in their instructional performance.

   Strongly Disagree   Strongly Agree

7) I believe our new teacher evaluation process is an effective tool to help teachers improve their performance.

   Strongly Disagree   Strongly Agree

8) I believe a teacher's special training beyond what is required by the district and continuing education should be considered in the evaluation process.

   Strongly Disagree   Strongly Agree

9) I believe a teacher's professional growth plan is an effective tool to improve his or her professional practice.
   Strongly Disagree   Strongly Agree

How many non-tenured classroom teachers did you personally evaluate during the 2011-2012 school year?

0   1-5   6-10   11-15 16-20 21-25 26-30 31-35 36-40 41-45

How many tenured classroom teachers did you personally evaluate during the 2011-2012 school year?

0   1-5   6-10   11-15 16-20 21-25 26-30 31-35 36-40 41-45

Does your school have a different evaluation process for tenure/non tenure teachers?

Yes   No

How likely were you to use the following evaluation instruments and procedures under your school's <u>former</u> (before Michigan's 2010/2011 teacher evaluation laws) teacher evaluation policy?
NA choice

SLIDING SCALE

	Very Unlikely	Very Likely
Pre-Observation Conference		
Classroom Walk Through		
Informal Conversations		
Evaluation Rubric		
Observation Checklist		
Post-Observation Conference		
Teacher Professional Growth Plan		
Teacher Portfolio		
End of Year Summative Narrative		

Under your school's <u>former</u> (before Michigan's 2010/2011 teacher evaluation laws) evaluation policy, on average how much time (hours) per year did you spend completing all aspects of teacher evaluation (preparation, observation process, meeting, write up, documentation, etc)?

_____ hours

How likely are you to use the following evaluation instruments and procedures under your school's <u>new</u> (after Michigan's 2010/2011 teacher evaluation laws) evaluation system?

SLIDING SCALE

	Very Unlikely	Very Likely
Pre-Observation Conference		
Classroom Walk Through		
Informal Conversations		
Evaluation Rubric		
Observation Checklist		
Post-Observation Conference		
Teacher Professional Growth Plan		
Teacher Portfolio		
End of Year Summative Narrative		

Under your school's <u>new</u> (after Michigan's 2010/2011 teacher evaluation laws) evaluation policy, on average how much time (hours) per year do you spend completing all aspects of teacher evaluation (preparation, observation process, meeting, write up, documentation, etc)?

_____ hours

REFERENCES

Blau, L. (2010, May 5). Overcoming the challenges of education policy implementation.

Retrieved from

www.kablaucommunications.com/wp-content/uploads/2011/02/Implementation-

FINAL.pdf

Boyd, R. (1989). Improving teacher evaluations. *Practical Assessment, Research*

And Evaluation, 1(7) 1-5.

Retrieved from http://PAREonline.net/getvn.asp?v=1&n=7

Center for American Progress. (2011). Access, evaluation, and effectiveness should guide

ESEA's teacher provisions.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, mixed methods*

approaches (3rd ed.). Thousand Oaks, CA: Sage

Daley, G., & Kim, L. (2010). A teacher evaluation system that works. National Institute for

Excellence in Teaching. Santa Monica, CA.

Danielson, C. (1996). *Teacher evaluation: To enhance professional practice: A framework*

for teaching. Alexandria, VA: Association for Supervision and Curriculum Development.

Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional*

practice. Alexandria, VA: Association for Supervision and Curriculum Development and

Educational Testing Service.

Darling-Hammond, L. (1990). Teacher evaluation in transition: Emerging roles and evolving

methods. In J. Millman & L. Darling-Hammond (Eds.) *The new teacher handbook of*

teacher evaluation: Assessing elementary and secondary school teachers. Newbury Park:

Sage.

Davis. D. R., Ellect, C. D. & Annunziata, J. (2002). Teacher evaluation, leadership and learning

organizations. *Journal of Personnel Evaluation in Education, 16*(4), 287-302.

Deal, T. E., & Peterson, K.D. (2009). *Shaping school culture.* San Fransisco, CA:

Jossey Bass.

Desimone, L. (2002). *How can comprehensive school reform models be successfully*

implemented? Retrieved from http://0-

search.proquest.com.catalog.lib.cmich.edu/docview/62216453?accountid=10181

Doherty, K. M., Jacobs, S., Lakis, K., Staresina, L., & Lasser, P. (2013). *Connect the dots:*

Using evaluation of teacher effectiveness to inform policy and practice. National Council

on Teacher Quality.

EdSource. (2009). *Glossary.* Retrieved April, 2011, from

http://www.edsource.org/pub_AccountabilityGlossary2-07.html

Education Nation. (NBC). (2012, September 24). Teacher town hall: Part 1. (Television

broadcast). New York, NY: MSNBC.

Ellet, C. D. & Teddlie, C. (2003). Teacher evaluation, teacher effectiveness, and school

effectiveness: Perspectives from the USA. *Journal of Personnel Evaluation in Education*,

17(1), 101-128.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed). London, UK: Sage.

Fisicaro, R. (2010). *Teacher evaluation: Assessing principals' perceptions in the state of*

New Jersey. Retrieved from ProQuest Digital Dissertations. (AAT 3391110)

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M. & Wallace, F. (2005).

Implementation research: A synthesis of the literature.

Fixsen, D. L., Blase, K. A., Duda, M. A., Naoom, & VanDkye, M. (2009). Making it happen:

Implementation drivers and stages. National Implementation Research Network.

Goe, L.G., Bell, C., & Little O. (2008). *Approaches to evaluating teacher effectiveness: A*

research synthesis. National Comprehensive Center for Teacher Quality.

Goe, L.G., Biggers, K., & Croft, A. (2012). *Linking teacher evaluation to professional*

development:Focusing on improving teaching and learning. National Comprehensive

Center for Teacher Quality.

Goldhammer, R. (1969). Clinical supervision: Special method for the supervision of

teachers. NY: Holt, Rinhart, and Winston.

Graham, M., Milanowski, A., & Jackson, M. (2012). *Measuring and promoting inter-rater*

agreement of teacher and principal performance ratings. Washington, DC: Center for

Education Compensation Reform. Retrieved from http://cecr.ed.gov/pdfs/Inter_Rater.pdf

Grossman, T. (2011). Preparing principals to evaluate teachers. *NGA center for best*

practice. Washington, DC. Retrieved from

http://www.danielsongroup.org/ckeditor/ckfinder/userfiles/files/NGA%20Issue-Brief-

Preparing-Principals-to-Evaluate-Teachers.pdf

Halverson, R., Kelley, C., & Kimball, S. (2004). Implementing teacher evaluation systems: How

    principals make sense of complex artifacts to shape local instructional practice. *Research*

    and theory in educational administration, 3. Retrieved from

    http://dm.education.wisc.edu/rrhalverson/intellcont/%20RTEA%20halversonkelleykimba

    ll%20110707-1.pdf

Hammond, L. D., Beardsley, A. A., Haertel, E. & Rothstein, J. (2012). Evaluating Teacher

    Evaluation: Popular modes of evaluating teachers are fraught with inaccuracies and

    inconsistencies, but the field has identified better approaches. *Phi Delta Kappan, 93*(6).

    Retrieved from

    http://www.edweek.org/ew/articles/2012/03/01/kappan_hammond.html

Harris, D. H. (2011). *Value-added measures in education: What every educator needs to know.*

    Cambridge, MA: Harvard Education Press.

Hazi, H, M., & Rucinski, D, A. (2009). Teacher evaluation as a policy target for

    improved student learning. *Education Policy Analysis Archives, 17*, 1-23.

Henry, Barton, S. N. (2010). *Principals' perceptions of teacher evaluation practices in an urban*

    school district. Retrieved from Pro Quest Digital Dissertations. (881470570).

Herman, J.L., Heritage, M., Goldschmidt, P. (2011). Developing and selecting assessments

    of student growth for use in educator evaluation systems. Assessment and Accoutability

    Comprehensive Center. Retrieved from

    http://www.cse.ucla.edu/products/policy/TestScoresTeacherEval.pdf

Ho, A.D., & Kane, T.J. (2013). The reliability of classroom observation by school personal. *Bill and melinda gates foundation.* Retrieved from http://www.metproject.org/downloads/MET_Reliability%20of%20Classroom%20Observations_Research%20Paper.pdf

Holtzapple, E. (2003). Criterion-related validity evidence for a standards-based teacher evaluation system. Journal of Personnel Evaluation in Education, 17(3), 207–219.

Honig, M.L. (2006). *New direction in education policy implementation: Confronting complexity.* Albany, NY: State University of New York Press.

Keesler, V., & Howe, C. (2012). Understanding educator evaluations in Michigan. Retrieved from http://www.michigan.gov/documents/mde/

Kersten, T. A., & Israel, M. S. (2005). Teacher evaluation: Principals' insights and suggestions for improvement. *Planning and Changing, 36*(1), 21-21. Retrieved from http://0-search.proquest.com.catalog.lib.cmich.edu/docview/62059839?accountid=10181

Lewin, K. (1951). *Field theory in social science: Selected theoretical papers.* New York, NY: Harper.

Little, O., Goe, L., & Bell, C. (2009). A practical guide to measuring teacher effectiveness. Retrieved from http://www.tqsource.org/publications/practicalGuide.pdf

Markley, T. (2004). Defining the effective teacher: Current arguments in education. *Essays in Education Journal, 11,* 1-14.

Marx, G.E. (2007). Teacher evaluation. The Principal's Partnership. Retrieved from http://educationpartnerships.org/pdfs/TeacherEvaluation.pdf

Marzano, R., Frontier, T., & Livingston, D. (2011). *Effective supervision: Supporting the art and science of teaching.* Alexandria, VA: ASCD.

129

Metz, A. J., Blase, K., & Bowie, L. (2007). Implementing evidence-based practices: Six "drivers" of success. *Research-to-Results Child Trends, 29.* Retrieved from www.childtrends.org/files/child_trends 2007_10_01_RB_6successdrivers.pdf

Measures of Effective Teaching Project. (2012). Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains. Retrieved from http://metproject.org/downloads/MET_Gathering_Feedback_Practioner_Brief.pdf

MET Project. (2013). Ensuring fair and reliable measures of effective teaching: Culminating findings from the met project's three-year study. Retrieved from http://metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_ Practitioner_Brief.pdf

Michigan Council for Educator Effectiveness. (2012). *Interim progress report.* Retrieved from http://www.mcede.org/reports

Michigan Council for Educator Effectiveness (2013). *Final recommendations.* Retrieved from http://www.mcede.org/

Michigan Department of Education. (2010). *Michigan moving forward with reforms despite not being chosen for Race to the Top.* Retrieved from http://www.michigan.gov/mde/0,1607,7-140-37818_53083-241276--,00.htm

Michigan Department of Education. (2011). *Educator evaluation systems.* Retrieved from http://www.michigan.gov/mde/0,4615,7-140-22709_57992---,00.html

Michigan Department of Education, Office of Special Education. (2011). MiBLSi implementation drivers reference sheet. Retrieved from http://miblsi.cenmi.org

Milanowski, A. (2004). The relationship between teacher performance evaluation scores and student achievement: Evidence from Cincinnati. Peabody Journal of Education, 79(4), 33–53.

Milanowski, A. T., Kimball, S. M., & Odden, A. (2005). Teacher accountability measures and links to learning. In L. Stiefel, A. E. Schwartz, R. Rubenstein & J. Zabel (Eds.), Measuring school performance and efficiency: Implications for practice and research (pp. 137–162). Larchmont, NY: Eye on Education.

Newton, X., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010). Value-added modeling of teacher effectiveness: An exploration across models and contexts. *Education Policy Analysis Archives, 18* (23). Retrieved from http://epaa.asu.edu/ojs/article/view/810

Pallas, A. (2012). Teachers: Why they quit, and why we can't fire our way to excellence. *Huff Post Education.* Retrieved from http://www.huffingtonpost.com/2012/08/08/teachers-why-they-quit-an_n_1753891.html

Rand. (2012). Measuring teacher effectiveness: A resource for teachers, administrators, policymakers, and parents. Retrieved from http://www.rand.org/education/projects/measuring-teacher-effectiveness.html

Rury, J. L. (2009). *Education and social change: Contours in the history of American schooling.* (3rd ed.). New York, NY: Routledge.

Sartain, L., Stoelinga, S.R., Brown, E.R., Luppescu, S., Matsko, K.K., Miller, F.K.,…Glazer, D. (2011). Rethinking teacher evaluation in chicago: Lessons learned from classroom observation, principal-teacher conferences, and district implementation. Consortium on Chicago School Research. Chicago, IL.

Schmoker, M. (2012). Why complex teacher evaluations don't work. *Education Week, 32*(2).

    Retrieved from

    http://www.edweek.org/ew/articles/2012/08/29/02schmoker_ep.h32.html?tkn=OTWFwcr

    Nk0aandQC5apgP1bbaNGPYzbMi16C&intc=es

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition:

    Reframing and refocusing implementation research. *Review of Educational Research.*

    72(3), 387-431. doi: 10.3102/00346543072003387

Taylor, E. S., & Tyler, J. H. (2012). Can teacher evaluation improve teaching? Evidence of

    growth in the effectiveness of midcareer teachers. *Education Next, 12*(4). Retrieved from

    http://educationnext.org/can-teacher-evaluation-improve-teaching/

Teachers' Tenure Act 4, Mic. Com. Law. 38-1-2 (1937 & Supp. 2011).

The Education Trust-Midwest. (2011). Teacher evaluation in Michigan. Retrieved from

    http://www.edtrust.org/midwest/publication/teacher-evaluation-in-michigan

The Danielson Group. (2010). The framework for teaching. Retrieved from

    http://www.danielsongroup.org/article.aspx?page=frameworkforteaching

The New Teacher Project. (2010). Teacher evaluation 2.0. Retrieved from

    http://tntp.org/

The New Teacher Project. (2012). Teacher evaluation 2.0. Retrieved from

    http://tntp.org/

The White House. (2010). *Education.* Retrieved from

    http://www.whitehouse.gov/issues/education/

Toch, T. (2008). Fixing teacher evaluation. *Educational Leadership, 66*(2), 32-37.

Tracy, S. J. (1995). How historical concepts of supervision relate to supervisory practices today.

    *Clearing House, 68(*5), 320.

United States Department of Education. (2009). *American recovery and reinvestment*

    act. Retrieved from http://www.ed.gov/recovery

United States Department of Education. (2010). *A blueprint for reform.* Retrieved from

    http://www2.ed.gov/policy/elsec/leg/blueprint/blueprint.pdf

United States Department of Education. (2010). *Race to the top.* Retrieved from

    http://www2.ed.gov/programs/racetothetop/index.html

Wagner, T. (2008). *The global achievement gap*. New York, NY: Basic Books.

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect: Our

    national failure to acknowledge and act on differences in teacher effectiveness. Retrieved

    from http://widgeteffect.org/downloads/TheWidgetEffect.pdf

Westat, M. G. (2011). What is inter-rater agreement and how can designers of teacher

    evaluation systems maximize it? Retrieved from

    http:// www.cecr.ed.gov/

    researchSyntheses/34008_CECR_RS_Inter_Rater_measurement_508.pdf

Werner. A. (2004). *A guide to implementation research.* Washington, DC: The Urban Institute

    Press.

Zepeda, S. J. (2012). *Informal classroom observations on the go*: *Feedback, discussion, and*

    reflection. Larchmont, NY: Eye on Education.