Culminating Projects in Education Administration and Leadership   Department of Educational Leadership and Higher Education

10-2017

The Value of Standards Based Teacher Evaluation: Perceptions of Tenured Elementary Teachers in a Metropolitan School District

Bryan Bass
St. Cloud State University

Follow this and additional works at: https://repository.stcloudstate.edu/edad_etds

Part of the Educational Leadership Commons

The Value of Standards Based Teacher Evaluation: Perceptions of Elementary Tenured

Teachers in a Metropolitan School District

by

Bryan E. Bass

A Dissertation

Submitted to the Graduate Faculty of

St. Cloud State University

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Education in

Educational Administration and Leadership

October, 2017

Dissertation Committee:
Roger Worner, Chairperson
Janine Dahms-Walker
Kay Worner
Nicholas Miller

Abstract

The authorization of federal legislation including the 2001 No Child Left Behind (NCLB) Act (Public Law 107-110) and the United States Department of Education, 2009 Race to The Top (RTTT) (Pubic Law 111-5) federal grant stimulus spurred an increase in state accountability systems that focused on reform and innovation. Specifically, these federal statutes focused on a "renewed emphasis on state-level teacher evaluation policy" (Maslow & Kelley, 2012, p. 601). This enhanced focus on accountability for student achievement among teachers and principals has contributed to school districts attempting to design effective teacher evaluation systems that promote and encourage a culture of continuous improvement and growth for all teachers. Further, there is a strong need to identify for principals the most essential elements of teacher evaluation that impact teacher effectiveness.

In the Minnesota State Teacher Development, Evaluation and Peer Support Model Evaluation Report (2015), teachers are "decidedly split" on the usefulness of teaching standards to accurately assess and inform professional growth conversations. The teachers who reported negatively identified that the tool was limiting and encouraged "canned ways of reflecting" on their instructional practices. However, evaluators found that the teaching standards are useful and that they need more time and/or training to support professional growth conversations with teachers.

Callahan and Sadeghi (2015) identify that a teacher evaluation system should provide "timely and useful feedback" through an "accurate and consistent process" that "measures a teacher's strength and weaknesses". Consequently, the purpose of this study is to understand the extent to which teachers perceive standards-based teacher evaluation to be useful and accurate in measuring teacher effectiveness. Additionally, the study seeks to identify how

teacher conferencing and written feedback within standards-based teacher evaluation influences teacher's professional growth (development).

 The frequency results provided several important findings. Thirty-four or 100.0% of respondents strongly agreed or agreed that conferencing (face-to-face) was an important component of teacher evaluation. Thirty-three or 97.1% of respondents strongly agreed or agreed that written feedback was an important component of teacher evaluation. Nineteen or 57.6% of respondents strongly disagreed or disagreed with the statement that the process of teacher evaluation leads to improved student achievement at their school. Thirty or 90.9% of respondents strongly agreed or agreed that when completed, conferencing (face-to-face) was helpful in improving teaching and learning.

Acknowledgements

I want to start by thanking my family. My wife, Katherine, who is my best friend and biggest supporter. Thank you for creating the space for me to complete this journey. To my daughters, Estelle and Veronica, thank you for listening to me throughout this journey and acting like you were interested when I wanted to share my excitement. It meant a lot for me to be able to share this experience with you. To my brother, Patrick, thank you for your belief in me and sharing your brilliant friend, Pearl, to give me another doctoral perspective in the final stages. To my mom and dad, thank you for your encouragement throughout my life to invest in education over sports. To my sister, Lisa, you are the most incredible role model. You are the exemplar for being resilient with whatever life throws at you.

Additionally, I want to thank my advisor, Dr. Roger Worner, for his unwavering support and belief in my ability to complete this journey. As a multiracial male, it means everything to me to have had your guidance and mentorship. Thank you for caring enough about my growth as a person that you demanded my best effort in completing the dissertation.

I want to also acknowledge my committee members, Dr. Janine Dahms-Walker, Dr. Kay Worner and Dr. Nick Miller, thank you for your enthusiasm and encouragement throughout the process. I want to also acknowledge Michele Braun-Heurung of the Center for Doctoral Studies at St. Cloud State University, for her timely support of me as a candidate. Michele was prompt and encouraging throughout the journey when I came to her with questions and concerns about next steps.

Lastly, I want to thank Al Levin, Mindy McBride, and Lisa Sayles-Adams, my colleagues and dear friends, who provided professional space and time for me to complete this experience.

Table of Contents

List of Tables

Chapter I: Introduction

Introduction

Historically, teacher evaluations systems in the United States have been criticized for their inability to achieve their intended purpose. Donaldson (2009) stated that teacher evaluation has not substantially improved instruction or expanded student learning. Toch (2008) identified that most school districts lack a credible system of measuring the quality of teachers' work. Maslow and Kelley (2012) asserted that the quality of evaluation suffers from school leaders' inconsistent implementation, competing demands and lack of clear understanding of how to assess high-quality teaching. Danielson (2011) argued that the purpose of teacher evaluation is "to ensure teacher quality" and "to promote professional development". Marzano (2012) identified that the two purposes of teacher evaluation were to measure teacher quality and develop a highly skilled workforce.

The authorization of federal legislation including the 2001 No Child Left Behind (NCLB) Act (Public Law 107-110) and the United States Department of Education, 2009 Race to The Top (RTTT) (Pubic Law 111-5) federal grant stimulus spurred an increase in state accountability systems that focused on reform and innovation. Specifically, these federal statutes focused on a "renewed emphasis on state-level teacher evaluation policy" (Maslow & Kelley, 2012, p. 601). Further, in 2009, Arne Duncan, Secretary of Education proposed that school districts report the percentage of teachers rated in each evaluation performance category (Donaldson, 2009, p. 1). Thus, there was an increased emphasis on accountability in the newly implemented teacher evaluation policy; focusing on measuring teachers' effectiveness toward increasing student achievement. Over 20 states have passed legislation changing teacher evaluation systems to reflect a greater emphasis on evidence of teachers'

impact on student achievement (Mead, Rotherham, & Brown, 2012). Recently, the state of

Minnesota enacted legislation that took effect in 2014 reflecting a similar emphasis.

In the fall of 2014, the State of Minnesota adopted The Teacher Development,

Evaluation & Peer Support Model in response to state legislative mandates (Minnesota State

Statute 122A.40, Subdivision 8 Development, evaluations, and peer coaching for continuing

contract teachers). The model featured three components including the following: 1) teacher

practice, 2) student engagement, and 3) student learning and achievement. The piloting of the

model began in the 2013-2014 school year, and a final report was completed by researchers

from the Center for Applied Research and Educational Improvement (CAREI) at the

University of Minnesota. In the report, teacher perceptions were collected through interviews

and revealed that the Minnesota Standards for Teacher Practice were "decidedly split between

those who found it useful and those who did not." However, evaluators viewed the standards

as useful and that they should be continued (Minnesota State Teacher Development,

Evaluation, and Peer Support Model Evaluation Report, 2015, p. 9). Evaluators further

reported that they needed additional time and training to provide effective feedback and

"…stimulate professional growth conversations."

This enhanced focus on accountability for student achievement among teachers and

principals has contributed to school districts attempting to design effective teacher evaluation

systems that promote and encourage a culture of continuous improvement and growth for all

teachers. Further, there is a strong need to identify for principals the most essential elements

of teacher evaluation that impact teacher effectiveness.

Research has revealed a breadth of standards recommended for use in evaluating

teacher effectiveness (Danielson's SOEI; InTASC; NBPTS). Nonetheless, there is limited

research that examined the perspectives of Minnesota metropolitan elementary school teachers on their experiences with a standards-based teacher evaluation process and whether or not they perceive it has improved teacher effectiveness.

Statement of the Problem

Through a comprehensive review of research on teacher evaluation, there appeared to be a gap in the literature that identified the effects of standards-based teacher evaluation or classroom observations on teacher effectiveness. The research indicated that teacher evaluation has historically and continues to be a perfunctory exercise that is significantly flawed (Tyler, Taylor, Kane, & Wooten, 2010; Donaldson, 2009; Sergiovanni and Starrat, 2007). Donaldson (2009) argued that the potential consequences, negative or positive, of teacher evaluation do not contribute to teachers' self-motivation to regard feedback that they receive from evaluation.

In the Minnesota State Teacher Development, Evaluation and Peer Support Model Evaluation Report (2015), teachers are "decidedly split" on the usefulness of teaching standards to accurately assess and inform professional growth conversations. The teachers who reported negatively identified that the tool was limiting and encouraged "canned ways of reflecting" on their instructional practices. However, evaluators found that the teaching standards are useful and that they need more time and/or training to support professional growth conversations with teachers.

The results of this preliminary report on the Minnesota State Teacher Development, Evaluation and Peer Support Model provided a limited snapshot of the teacher and evaluator (licensed administrator) perceptions on the usefulness of the model. As a result, in the state of Minnesota there is limited research on metropolitan elementary teacher perceptions on the

usefulness and accuracy of standards-based teacher evaluation tools. The study seeks to collect data from teachers on the usefulness and accuracy of standards-based teacher evaluation in advancing their teacher effectiveness.

Purpose of the Study

Maslow and Kelley (2012) maintain that teacher evaluation should ideally provide "meaningful feedback to teachers to improve teacher practice" and to be an "important source of data to inform organizational systems that support teaching and learning". Callahan and Sadeghi (2015) identify that a teacher evaluation system should provide "timely and useful feedback" through an "accurate and consistent process" that "measures a teacher's strength and weaknesses". Consequently, the purpose of this study is to understand the extent to which teachers perceive standards-based teacher evaluation to be useful and accurate in measuring teacher effectiveness. Additionally, the study seeks to identify how teacher conferencing and written feedback within standards-based teacher evaluation influences teacher's professional growth (development).

Research Questions

1. How did a select sample of tenured elementary school teachers from a Minnesota metropolitan school district differ in their perceptions of the accuracy of standards-based teacher evaluation in measuring teacher effectiveness?

2. What did a select sample of tenured elementary school teachers in a Minnesota metropolitan school district identify as useful components of standards-based teacher evaluation system?

3. What did a select sample of tenured elementary school teachers in a Minnesota metropolitan school district identify as those outcome(s) of their standards-based

teacher evaluation (including teacher conferencing and written feedback) were

most beneficial to them in improving their teacher effectiveness?

4. In what ways, did a select sample of tenured elementary school teachers in a

Minnesota metropolitan school district perceive written feedback from standards-

based teacher evaluation to impact professional growth?

Delimitations

According to Simon (2011) delimitations of the study defined the boundaries and limit

the scope of the research. Delimitations are within the researcher's control. The following

delimitations governed the scope and boundaries of this study.

The study's sample group was comprised of those teachers from a Minnesota

metropolitan school district. The sample does not include secondary school teachers.

The study's sample group was comprised of those elementary teachers who had

participated in a standards-based evaluation (including written feedback) component of a

teacher evaluation system in the selected Minnesota metropolitan school district. There are

demographic differences in the proportionality of female to male elementary teachers and

administrators.

Definitions of the Terms

Focus Rating: The State of Minnesota's measurement for identifying Focus Schools.

The Focus Rating is generated by combining the proficiency and growth of the seven

subgroups for which there is an achievement gap (Black, Hispanic, Asian, American Indian,

Free/Reduced Price Lunch, Special Education, and English Learners) (Minnesota Department

of Education, 2017).

Multiple Measurement Rating (MMR): The State of Minnesota's instrument for measuring a school's performance. The MMR measures proficiency, student growth, achievement gap reduction, and graduation rates. Schools earn points in each category based on student achievement. The percentage of points earned across the categories are totaled to determine a school's MMR (Minnesota Department of Education, 2017).

Probationary Teacher: Subdivision. 2. **Probationary period; discharge or demotion:** (a) All teachers in the public schools in cities of the first class during the first three years of consecutive employment shall be deemed to be in a probationary period of employment during which period any annual contract with any teacher may, or may not, be renewed as the school board, after consulting with the peer review committee charged with evaluating the probationary teachers under subdivision 3, shall see fit (Minnesota Statute: 122A.41 Teacher Tenure Act; Cities of the First Class; Definitions, 2015).

Principal: A school administrator who shapes the experiences of teachers and students through multiple interrelated roles, including building manager, employer, professional figurehead, supervisor, inspirational leader and provider of professional development (Leithwood & Louis, 2011; Waters, Marzano, & McNulty, 2003).

Professional Development: A comprehensive, sustained and intensive approach to improving teachers' and principals' effectiveness in raising student achievement (National Staff Development Council, 2016).

Standards-Based Teacher Evaluation: An evaluation of teacher performance that is based on a comparison with a set of standards that define effective teaching (Heneman, Milanowski, Kimball & Odden, 2006, p. 1).

Teacher Effectiveness: The impact that classroom factors, such as teaching methods, teacher expectations, classroom organization, and use of classroom resources, have on students' performance (Campbell, Kyriakides, Muijes, & Robinson, 2004, p. 3).

Tenured Teacher: a teacher who has received a continuing contract after successfully serving three consecutive years in a school district (Goldstein as cited in Kahlenberg, American Educator, 2015).

Summary

The perceptions of elementary tenured teachers in a Minnesota metropolitan school district on the usefulness and accuracy of a standards-based teacher evaluation system were examined in this study. The study is organized into five chapters, a bibliography, and appendices. The study opens with an introduction of the research problem. Next, a review of the related literature is examined through three themes. The themes included are measuring teacher effectiveness, standards-based teacher evaluation models and pivotal studies on teacher conferencing and written feedback.

The third chapter includes a description of the methodology used to collect data on the perceptions of elementary tenured teachers' standards-based teacher evaluation experiences. The results are detailed in chapter four. The fifth chapter provides a conclusion and recommendations for future study. The study concludes with a bibliography and appendices.

Chapter II: Literature Review

Introduction

This chapter will provide a review of the literature related to the research question of this study: What are elementary tenured-teachers' perception of standards-based teacher evaluation in a metropolitan district in the state of Minnesota? The researcher pursued this inquiry to uncover the instructional knowledge and understanding that teachers obtained from administrators using standards-based teacher evaluation system. Further, the researcher sought to reveal the effect that evaluation experiences had on the improvement of instructional practices school-wide (impact on teacher effectiveness). The first area of research provided a perspective on measuring teacher effectiveness—the emerging purpose of teacher evaluation outlined in federal and state legislation. The principle role for this section was to establish the increased focus on the importance of teacher evaluation in measuring teacher effectiveness.

The second area of research provided a review of the research related to the use of standards-based teacher evaluation models (methods). The purpose of this section was to identify the intended use and key components of standards-based teacher evaluation models. The final section of this chapter will review the research related to teacher conferencing and written feedback. The researcher sought to identify aspects of teacher evaluation that contribute to competent and effective evaluation experiences among teachers.

Measuring Teacher Effectiveness

Early research on assessing teacher effectiveness conducted by Kleinman (1966) acknowledged the gap or void in research on measuring teacher effectiveness. She reported, "Indeed, more than fifty years of research has not contributed much to our knowledge of factors which are associated with good teaching" (Kleinman, 1966, p. 234). Kleinman

identified early tensions with measuring teaching effectiveness that existed in the literature reviewed.

> One of the most difficult problems in the study of good teaching has been whether to assume that "effectiveness" is a statement about an attribute of a teacher in a particular teaching situation, or whether it is a statement about the results which come out of a teaching situation. (p. 234)

Kleinman (1966) concluded, "Measurement of behavior by observation appear to be the most promising technique to date for assessing teacher effectiveness" (p. 237). Thus, Kleinman's review of the research argued for the use of classroom observation instruments to measure teacher effectiveness.

Goe, Bell, and Little (2008) provided a more recent perspective on assessing teacher effectiveness in a study conducted on behalf of the National Comprehensive Center for Teacher Quality (TQ Center) in June 2008. First, they argued that the increase of school and classroom-level accountability is due, in part, to state and federal legislation (p. 2)—citing No Child Left Behind Act (NCLB, Public Law 107-110) and Race To The Top (RTTT, Public Law 111-115). The study sought to accomplish the following: "help regional and state decision makers better understand what constitutes effective teaching and the advantages and disadvantages of the various measures commonly used to evaluate it" (Goer et al., 2008, p. 3). Goe et al. (2008) argued that the definition needs to be comprehensive in view.

> Resist pressures to reduce the definition of teacher effectiveness to a single score obtained with an observation instrument or through using a value added model. Although it may be convenient to adopt a single measure of teacher effectiveness, there is no single measure that captures everything that a teacher contributes to educational, social and behavioral growth of students, not to mention ways teachers impact classrooms, colleagues, schools and communities. (p. 52)

A major theme throughout the study was the notion that assessing teacher effectiveness needs to be comprehensive. This was observed in their "five-point definition of effective teachers" (p. 8).

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

This definition focused on measuring multiple components of what Goe et al. (2008) argued meets a more comprehensive view of teacher effectiveness. Further, the definition "clarifies priorities" for measuring teacher effectiveness. They reported, "what is measured is a reflection of what is valued, and as a corollary, what is measured is valued" (p. 4).

Their study also included a review of current teacher evaluation measurement tools used in the field, including the following: a) Classroom Observation, b) Principal Evaluation,

c) Instructional Artifact, d) Portfolio, e) Teacher Self-Report Measure, f) Student Survey, and

g) Value-Added Model. This document is identified as *Table 1. Brief Summaries of Teacher*

Evaluation Methods (see Appendix A). The review of the teacher evaluation measurement

tools substantiated the claim that Goe et al. (2008) made with respect to the need for a

comprehensive view of defining teacher effectiveness.

> Seek other measures, or create appropriate measures, to capture important information about teachers' contributions that go beyond student achievement score gains. This may mean developing a measure that captures evidence of an individual teacher's leadership activities within the school, his or her collaboration with other teachers to strategize ways to help students who are at risk for failure, or participation in a study group to align curriculum with state standards. (p. 48)

This recommendation to maintain a comprehensive view in measuring teacher effectiveness

posed by Goe et al. (2008) was further researched by the Bill and Melinda Gates Foundation

in the Measure of Effective Teaching (MET) study in 2013.

Research from the Measures of Effective Teaching (MET) study (Bill & Melinda

Gates Foundation, 2013) identified three metrics in combination for determining teacher

effectiveness—Value Added Measure (VAM) scores or standardized test scores, classroom

observation instruments, and student perception surveys. The purpose for this study was to

provide "trustworthy information about teaching effectiveness" to support states and districts

in meeting the school-level and classroom-level requirements for measuring teacher

effectiveness as identified in RTTT (p. 3).

> States and districts have launched unprecedented efforts in recent years to build new feedback and evaluation systems that support teacher growth and development. The goal is to improve practice so that teachers can better help their students graduate from high school ready to succeed in college and beyond,\. (MET Project, 2013, p. 3)

Toward this end, the study used random assignments of teachers to students to determine if

measures of effective teaching identify teachers who enable student learning. They reported,

By definition, teaching is effective when it enables student learning. But identifying effective teaching is complicated by the fact that teachers often have very different students. Students start the year with different achievement levels and different needs. Moreover, some teachers tend to get particular types of students year after year (that is, they tend to get higher-performing or lower-performing ones). This is why so-called value-added measures attempt to account for differences in the measureable characteristics of a teacher's students, such as prior test scores and poverty. (p. 6)

The MET study found that by using multiple measures (student achievement gains, classroom observation instruments, and student perception surveys) that they were able to identify effective teachers. Further, the study analyzed various weights among the metrics to determine reliability on the impact of weighted metrics. The findings revealed that heavily weighting one measure over others is not predictive and limits a comprehensive view of teacher effectiveness. The study reported,

Heavily weighing a single measure may incentivize teachers to focus too narrowly on a single aspect of effective teaching and neglect its other important aspects…If the goal is for students to meet a broader set of learning objectives than are measured by a state's tests, then too-heavily weighting that test could make it harder to identify teachers who are producing other valued outcomes. (MET Project, 2013, p. 11)

The MET study (2013) concluded with the implications and recommendations for the implementation of the three combined metrics reviewed.

Standards-Based Teacher Evaluation Models

The emerging research on standards-based teacher evaluation models revealed disparate perspectives on the teacher evaluation process. First, it offered direct criticism of early iterations that employed limited and narrowly focused checklists for assessing instruction. Second, it demanded a more comprehensive view of effective teaching that reflects the breadth and depth of teaching. Nolan and Hoover (2008) asserted that the teacher evaluation process should not be reduced to a single check-list, but rather include a variety of effective teaching practices that appreciates the complexity of instruction. Sergiovanni and

Starratt (2007) affirmed this perspective on a checklist approach, "…this results in uniform use of an instrument that might be appropriate for a limited range of teaching and learning outcomes but may be invalid for other teaching and learning outcomes" (p. 172). Danielson and McGreal (2000) framed the demand for standards-based teacher evaluation similarly and add the need for the criteria to reflect current research on effective teaching. They reported, "The evaluative criteria used should represent the most current research available; and we need to make provisions, as time goes on, to revise those criteria to reflect current findings" (Danielson & McGreal, 2000, p. 3).

Toward this end, Nolan and Hoover (2008) provided a description of standards-based teacher evaluation.

> In contrast to checklists that rate teachers on the basis of their fidelity to particular models of instruction, a standards-based approach to effective teaching specifies a given number of capabilities, applicable in a variety of contexts, that teachers are expected to possess. For each standard, stated in broad terms, there are multiple indicators that can be identified in a teacher's performance to show evidence that the teacher possesses that capability. These indicators, in contrast to specific behaviors required by a model of teaching, may be demonstrated using a variety of teaching models or approaches that are appropriate for the goals of the lesson, the subject, and the particular group of learners. (p. 174)

This description emphasized the context of teaching as critical in developing standards that are relevant and applicable for teachers. Flanders (1976, as cited in Nolan & Hoover, 2008) argued that context matters in teacher evaluation. Further, Nolan and Hoover (2008) argued the development of standards-based teacher evaluation models at the local level include key stakeholders. They reported, "Development of standards of effective teaching locally, as we define it here, may take the form of either creating the district's own set of standards or modifying standards developed by external sources" (Nolan & Hoover, 2008, p. 174). Danielson and McGreal (2000) agreed, "…in designing (or revising) its system of evaluation,

a school district should follow a process that includes many perspectives—those of teachers, administrators, and the leadership of the teacher's association (p. 21).

Nolan and Hoover (2008) maintained that in addition to key experts who have developed teacher evaluation standards that there are "hundreds or thousands of teaching standards in existence" (p. 174). For the purposes of this research study, the researcher will present the teaching standards established in Charlotte Danielson's Framework for Teaching (2013) and the National Board for Professional Teaching Standards (NBPTS) (2015).

Danielson's framework integrated the research (and frameworks) from the Praxis III tool for assessing teacher candidates for initial licensure, the INTASC (Interstate New Teacher Assessment and Support Consortium) and NBPTS. Danielson's framework included four domains and 22 teaching components. Levels of performance for each teaching component are configured in a teacher behavior rubric (see Appendix). The teacher behavior rubric includes four levels of performance: unsatisfactory, basic, proficient, and distinguished. Table 2.1 lists the domains and components (standards). Danielson (1996), in her original edition of "A Framework for Teaching", emphasized that the domains and components reflect the complexity of teaching.

> Because teaching is complex, it is helpful to have a road map through the territory, structured around a shared understanding of teaching. Novice teachers, of necessity, are concerned with day-to-day survival; experienced teachers want to improve their effectiveness and help their colleagues do so as well; highly accomplished teachers want to move toward advanced certification and serve as a resource to less-experienced colleagues. (Danielson, 1996, p. 2)

Nolan and Hoover (2008) in a review of Danielson's framework, maintained that the standards included in her framework provide "a variety of approaches to data gathering in addition to focused classroom observations" (Nolan & Hoover, 2008, p. 175). Sergiovanni

and Starratt (2007) argued that this dimension of the framework is key as earlier attempts at

assessing instruction were criticized as too narrow or merely a script.

> Rather than being a script, the framework helps teachers by forcing them to consider
> many possibilities and to discuss with their supervisors what makes sense in a
> particular situation, as well as how the various components might look in practice as
> situations vary. (Sergiovanni & Starratt, 2007, p. 186)

Table 2.1

Charlotte Danielson's (2013) Domains and Standards

Domain	Component
1. Planning and Preparation	1a. Demonstrating knowledge of content and pedagogy 1b. Demonstrating knowledge of students 1c. Setting instructional outcomes 1d. Demonstrating knowledge of resources 1e. Designing coherent instruction 1f. Designing student assessments
2. The Classroom Environment	2a. Creating an environment of respect and rapport 2b. Establishing a culture for learning 2c. Managing classroom procedures 2d. Managing student behavior 2e. Organizing physical space
3. Instruction	3a. Communicating with students 3b. Using questioning and discussion techniques 3c. Engaging students in learning 3d. Using assessment in instruction 3e. Demonstrating flexibility and responsiveness
4. Professional Responsibilities	4a. Reflection on teaching 4b. Maintaining accurate records 4c. Communicating with families 4d. Participating in a professional community 4e. Growing and developing professionally 4f. Showing professionalism

Danielson has since revised the framework three additional times including 2007, 2011, and

2013. Each time attempting to tighten the language, broaden its practical application, increase

its relevance for members in non-classroom specialist positions (counselors, librarians, and

nurses), and reflect the current research informing what constitutes effective teaching.

Danielson clarified these distinctions in the 2013 edition. Danielson (2013) noted several

influences that impacted the revisions. In 2009, the framework was selected by the Bill and

Melinda Gates Foundation's pivotal research project entitled the Measures of Effective

Teaching study also known as the MET study. The MET study sought to "determine which

aspects of teacher's practice were most highly correlated with high levels of student progress"

(Danielson, 2013, p. 1). This large scale study included the video capture of over 23,000

lessons. The outcome of the involvement in this study led to significant enhancements in the

revision of the tool in 2011.

> These enhancements to the Framework for Teaching, while created in response to the
> demands of the MET study, turned out to be valuable additions to the instrument in all
> its applications. Practitioners found that the enhancements not only made it easier to
> determine the level of performance reflected in a classroom for each component of the
> Framework, but also contributed to judgments that are more accurate and more worthy
> of confidence. As the stakes in teacher evaluation become higher, this increased
> accuracy is absolutely essential. (Danielson, 2013, p. 2)

The second substantial influence to the final revision of the Framework for Teaching

(which occurred in 2013) was the adoption of the Common Core State Standards (CCSS) in

most states. Danielson (2013) argued that the adoption of the CCSS demands teachers to

"acquire new instructional skills" in order to effectively teach "deep conceptual

understanding, for argumentation and for logical reasoning" (Danielson, 2013, p. 3). In the

2013 edition, Danielson concentrated the integration of CCSS philosophy into the "Possible

Examples" section for each level of performance. She reported, "…many of the enhancements

to the Framework are located in the possible examples, rather than in the rubric language or

critical attributes for each level of performance" (Danielson, 2013, p. 3).

It has been noted in the research (Nolan & Hoover, 2008; Sergiovanni & Starratt,

2007) that the use of the Framework for Teaching has been piloted in teacher evaluation

systems developed in local districts across many states. Heneman et al. (2006) studied the use

of the framework as a key component in a teacher evaluation system in four sites throughout

the country.

> The Framework for Teaching (with adaptation to the local context) can be used as the performance measure for a standards-based teacher evaluation system. Evaluators can gather evidence from various sources (e.g., classroom observation, portfolios, logs) about the teacher's performance and then rate the teacher's performance on each element. Written and verbal feedback can be provided, and action plans for improvement can be developed. (Heneman et al., 2006, p. 2)

In their study, Heneman et al. (2006) provided a visual that depicts the adaptations of the

Framework for Teaching made by the four sites included in their study (see Appendix). The

visual demonstrates various uses of Danielson's framework, ranging from sites including *all*

of the domains and components to *some*. This reveals that Danielson's intended use of the

tool to stimulate professional conversations about effective teaching at the local level is

occurring. The study conducted by Heneman et al. (2006) sought to find answers to the

following questions related to the use of Danielson's framework as a competency model for

teacher evaluation.

- What is the relationship between teachers' standards-based teacher evaluation scores or ratings and the achievement of their students?

- How do teachers and administrators react to standards-based teacher evaluation as a measure of instructional expertise?

- Is there evidence that standards-based teacher evaluation systems influence teacher practice?

- Do design and implementation processes make a difference?

Their research findings were based on quantitative and qualitative research methods. The

authors established a value-added formula that included a 3-year average of academic

achievement scores for each teacher to determine a correlation effect to evaluation ratings or

scores. They interviewed teachers and evaluators and conducted multiple surveys to

understand teacher and administrator reactions as well as the impact of the framework on

teacher practice. Two districts stood out in their training and preparation for teachers and

evaluators and it contributed to a positive correlation between teacher ratings and students'

academic achievement.

> We speculate that Cincinnati and Vaughn have higher average correlations in part due to the use of multiple evaluators. In addition, Cincinnati evaluators received intensive, high-quality training. Vaughn evaluators could draw on a strong shared culture and history of working on instruction that fostered agreement on what good teaching looks like. (Heneman et al., 2006, p. 5)

Teacher and administrator reactions were "most positive and least varied" when responding to

the performance competency model (Danielson's framework adapted). It became clear in the

findings that the opportunity to have a model that established a common language for talking

about effective teaching was essential to the teachers and evaluators.

> Many teachers told us that this was the first time they ever had a clear and concise understanding of the district's performance expectations for their instructional practice. Additionally, many reported that the use of the teaching standards helped improve dialogue with their principals about teaching and performance expectations…Many principals valued the increased opportunity to discuss instruction with teachers and felt that the greater the amount of evidence they collected, combined with the explicit rubrics describing the four levels of teacher performance, helped them do a better job as evaluators. (Heneman et al., 2006, p. 6)

The impact on teacher practice was "broad, but relatively shallow" (Heneman et al., 2006, p.

7). Identification of positive impacts were identified as increased reflection, improved lesson

planning, and better classroom management.

Heneman et al. (2006) also revealed that at all sites in their study, "administrator

training did not appear to put much emphasis on providing useable feedback, setting

performance goals and coaching" (Heneman et al., 2006, p. 7). This finding will be further explored in the next section of this chapter when examining teacher feedback from observations.

The National Board for Professional Teaching Standards has established a set of standards to support excellent teaching where teachers seek to become board certified. Sergiovanni and Starratt (2007) argued that "Board certified teachers are considered among the nation's best" (p. 194). They identified that this process is rigorous.

> The process provides an opportunity for teams of teachers to work together in discussing and understanding the board standards and to help each other prepare the necessary documents required by the assessment process. And the process stretches those who are principals or other designated supervisors by increasing their own learning curves and calling on them to place teaching and learning at the center of their practice. (Sergiovanni & Starratt, 2007, p. 194)

NBPTS identified five core propositions and their seventeen indicators for accomplished teaching, Table 2.2 lists them. The standards were recently revised in the 2016 edition of "What Teachers Should Know and Be Able to Do". The authors, on the outset, identified the most critical aspect of the revisions is that they were completed by and for teachers—a staple of NBPTS.

> A distinguishing hallmark of a profession is that those who are in it determine what its members must know and do. For this reason, how these revisions took place is as important as the revisions themselves. As is the case with all National Board Standards, the updated Five Core Propositions were written by teachers, for teachers. The Five Core Propositions—in content and in authorship—area a statement of what our profession stands for. (NBPTS, 2016)

Similar to the process recommended by Danielson in the Framework for Teaching, teachers in conversations with teachers about what constitutes effective teaching results in a more applicable instrument. The outcome of this demanding certification is that teachers know

they're being measured against criteria that were developed by their peers, rather than

government institutions.

Table 2.2

National Board for Professional Teaching Standards (revised 2016)

Five Propositions	Indicators
1. Teachers are committed to students and their learning.	A. Teachers recognize individual differences in their students and adjust their practice accordingly. B. Teachers understand how students develop and learn. C. Teachers treat students equitably. D. Teachers know their mission transcends the cognitive development of their students.
2. Teachers know the subjects they teach and how to teach those subjects to students.	A. Teachers appreciate how knowledge in their subjects is created, organized and linked to other disciplines. B. Teachers command specialized knowledge of how to convey a subject to students. C. Teachers generate multiple paths to knowledge.
3. Teachers are responsible for managing and monitoring student learning.	A. Teachers call on multiple methods to meet their instructional goals. B. Teachers support student learning in varied settings and groups. C. Teachers value student engagement. D. Teachers regularly assess student progress. E. Teachers engage students in the learning process.
4. Teachers think systematically about their practice and learn from experience.	A. Teachers make difficult choices that test their professional judgment. B. Teachers use feedback and research to improve their practice and positively impact student learning.
5. Teachers are members of learning communities.	A. Teachers collaborate with other professionals to improve school effectiveness. B. Teachers work collaboratively with families C. Teachers work collaboratively with the community

Teacher Conferencing and Feedback

In this section, the researcher seeks to identify in the related literature, the influence of

teacher conferencing and feedback on improving teacher effectiveness.

The research on teacher evaluation identifies two fundamental purposes of teacher

evaluation: "quality assurance and professional learning" (Danielson, 2007, p. 64). Haefele

(1993) as cited in Feeney (2007) argued that evaluation should support the goals of screening

out unqualified teachers, maintaining high level of quality instruction, and providing

constructive feedback to support teachers' professional growth. Frase (2001) completed a

study that revealed constructive feedback is missing in teacher evaluation. He cited state

audits (from 1990) of teacher evaluations and curriculum from five school districts in North

Carolina as evidence that evaluation scores are inflated and little feedback for improvement

are provided by evaluators. Frase reported, "In most districts, no teachers, including

probationary teachers, were found to be below standard, while the large majority were rated

above standard" (2001, p. 177). Frase also reported, "Auditors' observations revealed poor

instructional practices, particularly in districts where evaluation ratings were highest" (2001,

p. 178). Frase (2001) claimed that constructive feedback is not being provided or is

completely missing from teacher evaluation.

> Teachers in many cases have good reason for holding evaluation and supervision in contempt. Evaluations have not been helpful; evaluators are not trained in curriculum and instruction, and feedback is either absent or of low quality. Most serious of all, they do not result in instructional improvement. Failure to provide accurate feedback accompanied by substantive and practical suggestions for improvement closes the door to improvement and enhanced intrinsic motivation. (Frase, 2001, p. 178)

Feeney (2007) agreed with Frase's criticism on the lack of constructive feedback. He

reported, "Without quality feedback to inform teaching, a teacher's independent creation of

meaningful goals for his or her own professional growth probably will not happen" (Feeney,

2007, p. 192). Feeney (2007) identified three criteria for effective or quality feedback based

on a review of the literature.

1. Be based on descriptive observable data (Danielson & McGreal, 2000);

2. Provide characteristics of effective teaching (Danielson, 1996; Marzano,

   Pickering, & Pollock 2001);

3. Promote reflective inquiry and self-directedness to foster improvements in

   teaching supported by evidence of student learning (Glickman, 2002).

Feeney (2007) completed a case study to determine if the three criteria identified above would be met by the use of a performance rubric (adapted from Danielson's Framework for Teaching, 1996) by evaluators when providing feedback. Feeney (2007) used narrative feedback from evaluations written by several different administrators from 1982 to 2006 (confirm these dates—does not jive with post 1999 comparison). Feeney (2007) acknowledges that a performance rubric was adopted or implemented in 1999. Therefore, when reviewing the data prior to 1999, he discovered a pattern of feedback that "did not promote and support professional learning" and "was nondescript, meaningless at times, and did not reference student-learning outcomes" (Feeney, 2007, p. 193). When looking at the data after 1999, Feeney reported, "A performance rubric on effective characteristics of teaching provides a focus for evaluators to use when providing direction to teachers throughout the evaluation process" (Feeney, 2007, p. 194). Feeney (2007) acknowledged that it is unclear if the use of the performance rubric was able to meet all three criteria.

> From this case study, it is unclear if the feedback in a summative evaluation promotes reflective inquiry and self-directedness in teachers to foster improvements in teaching supported by evidence from student learning. A review of administrator feedback using the criteria for effective feedback supports the claim that identifying characteristics of effective teaching helps evaluators provide quality constructive feedback to teachers. (Feeney, 2007, p. 194)

 A significant concern expressed by Feeney (2007) was the intrinsic motivation that the teacher has to routinely engage in reflective inquiry stemming from bidirectional communication throughout the evaluation process. The role of the administrator becomes critical in how he/she engages a systematic framework for teacher conferencing and feedback that allows all three criteria to be met.

Zimmerman and Deckert-Pelton (2003) studied teachers' perceptions of the principal in professional evaluation. The study examined the responses of 86 educators from five northwest Florida counties. Using a constant comparative method of analysis (qualitative), Zimmerman and Deckert-Pelton (2003) sought to collect responses to the following questions: 1) How do teachers view the principal as their primary evaluator? 2) What do they perceive is the principal's role in the evaluation process? and 3) What makes a principal an effective evaluator?

Zimmerman and Deckert-Pelton (2003) found that survey responses were focused in four domains: interaction, consistency, commitment and knowledge. In the area of interaction, teachers identified characteristics of effective evaluative interactions including the following: constructive general feedback, encouragement, pedagogically appropriate feedback, and adequate time for the feedback process. In the area of evaluative consistency, teachers reported a lack of consistency among principal ratings. One statement included to illustrate this point identified "what is expected as a good answer at one school is often not a good answer at another" (Survey respondent as cited in Zimmerman & Deckert-Pelton, 2003, p. 33). In the area of commitment, teachers reported that the principal's approach (or mindset) to teacher evaluation impacts the overall experience. In the last area of knowledge, Zimmerman and Deckert-Pelton (2003) found that teachers reported a sense of more confidence in principals that were formerly effective teachers are more effective evaluators.

McGreal (1983) as cited by Zimmerman and Deckert-Pelton (2003), reported, "the relationship between a principal and faculty members has a pivotal effect on instructional effectiveness" (p. 29). Zimmerman and Deckert-Pelton (2003) also cited Valentine (1992), where he maintained "that implementing the improvement component is one of the most

challenging tasks of the principal in the performance-based developmental evaluation process" (p. 29). Zimmerman and Deckert-Pelton (2003) concluded that the principal has a pivotal role in the effective use of teacher evaluation processes to improve teacher effectiveness.

> In conclusion, principals must carefully evaluate their own knowledge, skills, and abilities with regard to the critical process of teacher evaluation. They must be willing to adapt to the new expectations for today's educational systems and provide inspired, knowledgeable, and imaginative evaluations. (Zimmerman & Deckert-Pelton, 2003, p. 35)

Anast-May, Penick, Schroyer, and Howell (2011) completed a study on the perceptions of teachers' experiences with teacher conferencing and feedback as a systematic structure (component) of a standards-based teacher evaluation process. The research was conducted in a large South Carolinian school district including 37 teachers. The researchers' reported, "To provide quality feedback, a structure needs to occur to promote reflective inquiry and conversations for facilitating the learning of teachers" (Anast-May et al., 2011, p. 3). In additional, Anast-May, Penick, Schroyer, and Howell (2011) cited research that there is strong potential for bias in results of teacher evaluation when there is lack of frequent observations. Denner, Miller, Newsome and Birdsong (2002, as cited in Anast-May et al., 2011) maintained that when observations occur frequently, their reliability improves and when observations are longer, their validity improves (Cronin & Capie, 1986 as cited in Anast-May et al., 2011).

Anast-May et al. (2011) also acknowledged a concern about the extent to which feedback structures (conferencing) when combined with frequent observation actually "promote reflective inquiry and self-directedness to foster improvements in teaching supported by evidence of student learning" (Glickman, 2002 as cited in Anast-May et al.,

2011, p. 3). Therefore, their study sought to identify teacher perceptions of receiving face-to-face feedback in a conferencing structure. The observation process included three assistant principals representing different elementary schools performing five observations per teacher, 60 minutes per observation, and pre and post conferencing including descriptive observable data with feedback.

The findings were significant in the survey responses from the participants. Anast-May et al. (2011) reported, "When responding to the survey as to how face-to-face conferencing benefitted participants, all of the teachers who participated in the pre and/or post conferences felt that the conferences were positive and assisted them in their professional growth" (p. 5). A key finding was that the face-to-face conferencing had a powerful impact on teacher's ability to understand more than just the observable data in the written feedback.

All 37 participants responded that nothing takes the place of face-to-face conferencing nor can the same information be relayed in an email or a little note left on the desk after the observation. Face-to-face discussion is more powerful and has a greater impact (Anast-May et al., 2011, p. 5).

The survey responses collected by Anast-May et al. (2011) also acknowledged another key finding related to the frequency of observations that should be completed throughout the year. They reported, "As to how often the observations should occur, 17% responded bi-weekly, 49% responded monthly, 13% responded quarterly, and 21% responded as often as possible" (Anast-May et al., 2011, p. 5). This key finding in combination of the finding around the face-to-face conferencing is pivotal in the related literature. Feeney (2007) as cited earlier in this section, offered three criteria for quality feedback. This study suggested a systematic structure for providing quality feedback to teachers in combination with a

performance rubric (or locally agreed upon standards of effective teaching) that reported

significant confidence in the process by teachers.

Synthesis of the Review of the Research

This chapter provided a review of the literature related to the research question of this

study: What are elementary tenured-teachers' perception of standards-based teacher

evaluation in a metropolitan school district in the state of Minnesota?

The first section identified the pivotal study completed by the Bill and Melinda Gates

Foundation, Measures of Effective Teaching study. The MET study provided three critical

components of a teacher evaluation system that support the identification of effecting

teaching: standards-based observation tool, student perception surveys and student scores.

The second section reviewed the research completed by Sergiovanni and Starrat

(2007) and Nolan and Hoover (2008) on the emergence of standards-based evaluation tools.

Specifically, the examination of Charlotte Danielson's Framework for Teaching (1996, 2007,

2011, and 2013) and the National Board for Professional Teaching Standards (1999). The

research revealed the importance of developing teaching standards that reflect the complexity

of teaching. It is important to note that the continued revision of Danielson's framework

highlighted the increasing demands of diverse language learners and a commitment to

interrupting racial disparities in curriculum and instruction, reinforcing the view that

standards should reflect a comprehensive view of teaching (MET, 2013).

The final section reviewed the research on teacher conferencing and feedback. Frase

(2001), Zimmerman and Deckert-Pelton (2003), Feeney (2007), and Anast-May et al. (2011)

all provided significant studies reviewing factors that contribute to quality feedback. The

development of three criteria contributed by Feeney (2007) paired with a performance rubric

(locally designed or adapted from existing frameworks) and systematic structure for teacher conferencing (face-to-face) and feedback led to teachers reporting greater confidence in the teacher evaluation process contributing to their improved practice and reflective inquiry.

Summary

The literature review provided an examination of three themes in the field of standards-based teacher evaluation. They included measuring teacher effectiveness, standards-based teacher evaluation models, and the impact of teacher conferencing (face-to-face) and written feedback. In Chapter III, the next chapter, the researcher describes the methodology used to collect data on elementary tenured teachers' perceptions of standards-based teacher evaluation experiences.

Chapter III: Methodology

Introduction

Federal legislation including No Child Left Behind (NCLB, Public Law 107-110) and the Race to The Top (RTTT, Public Law 111-115) federal grant stimulus have provided impetus for reform and innovation in teacher evaluation throughout the United States, including in the state of Minnesota. This intensified focus on accountability for teachers and principals aimed at increasing student achievement has contributed to school districts attempting to design effective teacher evaluation systems that promote and encourage a culture of continuous improvement and growth for all teachers. Accompanying the design of effective teacher evaluation systems is the perceived need to question principals in their understanding of the most essential elements of teacher evaluation for increasing teacher effectiveness.

Research revealed a breadth of standards that have been recommended for use in evaluating teacher effectiveness (Danielson's SOEI; InTASC; NBPTS). However, there is limited research that reflected the perspectives of those Minnesota metropolitan elementary teachers who are experiencing a standards-based teacher evaluation process on whether or not the process has improved teacher effectiveness.

The study examined the perceptions of elementary teachers from a Minnesota metropolitan school district on the accuracy and usefulness of standards-based teacher evaluation on teacher effectiveness. An original framework developed by Jonathon D. Pizzi (2009) that included an examination of urban secondary school administrators and teachers' perceptions of standards-based teacher evaluation in an urban high school in Boston, Massachusetts was replicated.

The four questions explored in the study were as follows:

1. How did a select sample of elementary school teachers from a Minnesota metropolitan school district differ in their perceptions of the accuracy of standards-based teacher evaluation in measuring teacher effectiveness?

2. What did a select sample of elementary school teachers in a Minnesota metropolitan school district identify as useful components of a standards-based teacher evaluation system?

3. What did a select sample of elementary school teachers in a Minnesota metropolitan school district identify as those outcome(s) of their standards-based teacher evaluation (including teacher conferencing and written feedback) that were most beneficial to them in improving their teacher effectiveness?

4. How did a select sample of elementary school teachers in a Minnesota metropolitan school district perceive written feedback from standards-based teacher evaluation of value to them in impacting their professional growth?

Participants

The study sought data from elementary teachers in the state of Minnesota from a metropolitan school district in the spring of 2017. The population selected was based on the recommendations from an original study by Jonathon Pizzi in 2009. The Pizzi (2009) study revealed that the examination of the perceptions of only secondary teachers' limited the generalizability of the study. Pizzi (2009) recommended that future study should include a larger sample size with teachers from elementary, middle and senior high school. Therefore, the tenured-teacher sample population were identified in cooperation with the metropolitan school district's department of research, evaluation, and assessment. In April, 2017 email

addresses of all tenured teachers from seven randomly selected elementary schools were provided by the school district's department of research, evaluation and assessment.

Human Subject Approval—Institutional Review Board (IRB)

The researcher completed all required Institutional Review Board training modules to earn certification as the investigator of the study. The Institutional Review Board at St. Cloud State University reviewed the application for this research study and (to be completed upon final approval from the SCSU IRB).

Research Design

The researcher used mixed methods to collect data modeled after a study on secondary teachers' perceptions in Boston, Massachusetts (Pizzi, 2009). Data collection included two phases: 1) a focus group to revise the survey tool; and 2) dissemination of the survey tool utilizing a Likert attitude scale and open-ended response questions. The Likert attitude scale provided quantitative data. The open-ended response questions were qualitative data that allowed for clarifying responses from the Likert attitude scale.

Instruments for Data Collection and Analysis

A mixed methods survey tool was used to collect data for the study. The mixed methods survey instrument was replicated from Pizzi's 2009 Teacher Evaluation System Attitude Scale that was used to research secondary teachers' perceptions of a teacher evaluation system used in Boston Public Schools in Massachusetts. Pizzi's 2009 Teacher Evaluation System Attitude Scale included five sections that consisted of 21 items on a Likert scale, four open response questions, and the identification of professional demographic information. The 21 items reflected three aspects of teacher evaluation: 1) teaching standards, 2) evaluation activities and communication, and 3) evaluation and instructional improvement.

The four open response questions reflected a qualitative method of collecting input on teachers' perceptions on the effectiveness of teacher evaluation activities and their impact on improving teacher practice (Pizzi, 2009, p. 64).

For the purposes of this study, the researcher modified 7 of the 21 items, added one additional item to make a total of 22, two of the open-response questions and three of the professional demographic questions. Several modifications to the Pizzi (2009) tool were initially completed by the researcher to include teaching standards used in the identified metropolitan school district. The researcher modified "Section A: Items 1-6, The Seven Standards for Teacher Performance" to "Section A: Items 1-6, The Standards of Effective Teaching" (see Appendix C: The Pilot Survey). The "Standards of Effective Teaching" consists of the teaching standards and the performance rubric used in the formal observation cycles of teachers in the identified metropolitan school district. The researcher also modified "Section E: Items 26-31, Professional Demographic Information" to eliminate the identification of probationary teachers. The focus of this research sample is elementary tenured teachers' perceptions.

The researcher sought to strengthen the original wording of "Section B: Items 7-14, Evaluation Activities and Communication" to reflect the related literature included in Chapter II. Specifically, increasing the deliberate emphasis on teacher conferencing (face-to-face) and written feedback. The researcher achieved this intention by rephrasing the section to the following: "Section B: Items 7-14, Evaluation Activities, Conferencing and Written Feedback" (see Appendix C). The researcher also added "face-to-face" in parenthesis next to all uses of conferences within Section B. This adapted the tool to allow for respondents to report out the impact of conferencing on teacher evaluation experiences.

The researcher then tested the tool with a pilot group consisting of a convenience sample of six tenured elementary teachers from the identified metropolitan school district. The six classroom teachers selected ranged from Kindergarten through fifth grade. They were selected as to be reflective of multiple perspectives across the elementary grades. The survey was emailed to the six teachers. Five of the six respondents communicated their feedback via email. One respondent printed the survey out and forwarded handwritten feedback. The feedback provided related to the following areas: 1) demographic identification, 2) wording of select items related to teaching standards, evaluation activities and communication, and 3) the phrasing of two open response questions, 23 and 25.

Pilot group responses related to demographic identification included a concern for identifying gender on a broader continuum than just male or female as we may have teachers who identify as transgender. Additionally, the feedback included a concern for identifying the race of individuals who may be more than one race and who may not be represented in the categories provided. These concerns were reviewed and modified by the researcher in review of the guidance provided by the United States Department of Education in October, 2007 entitled Final Guidance on Maintaining, Collecting, and Reporting Racial and Ethnic Data to the U.S. Department of Education (U.S. Department of Education, 2007). The researcher also received guidance from St. Cloud State University around its current research practices for demographic identification for race and gender.

In addition, pilot group respondents provided technical feedback regarding the phrasing of "…in order to…" in items thirteen through fifteen to using "…to…". The researcher accepted the recommended phrasing changes and further modified the tool. The last area of feedback received was related to the wording of two open response questions. The

use of the word "substantial" in item 23, was identified as confusing by one of the

respondents. The recommendation by the respondent was to use the word "significant" in

place of "substantial". The researcher accepted this suggestion. The other wording

recommendation was in open response question 25 which included the words "warm and cool

feedback". The respondent recommended to eliminate "warm and cool" and simply use the

phrase "any additional feedback". The researcher applied all recommendations provided to

improve the potential effectiveness of the tool.

After all feedback was collected and applied from the pilot group, the final instrument

reflected a similar layout to Pizzi's (2009) tool including the modifications articulated above.

The five sections, included twenty-two items and four open response questions to maintain

reliability and validity of the instrument. The research is advancing the original study of Pizzi

(2009) and validated similarly with a convenience sample or pilot group and committee

review.

Procedures and Timeline

Data collection was initiated in mid-May and ended June, 2017. In early May the

researcher provided an introduction letter (see Appendix E) to the principals of the seven

randomly selected schools requesting their support in encouraging tenured teachers from their

school to participate in the study.

An email invitation was sent in mid-May to all selected tenured elementary teachers

and their assigned administrators (see Appendix F) through Survey Monkey.

A reminder email was sent 1 week after the initial invite to increase the response rates

of participants who had not yet responded. A final email was sent the last week of May to

increase the participation rate to over 70%.

Limitations

There are several limitations that are specific to the design of the study. First, the study

is limited by the sample size of the population of survey respondents. The researcher selected

one metropolitan school district to control for the variability in training of evaluators, the

process of teacher evaluation as proscribed by the school district, and knowledge of the

teacher evaluation process by the participants. Additionally, the researcher narrowed the

population to tenured elementary teachers. This limits the potential number of participants as

there are a finite number of tenured teachers.

The second limitation is related to the design of the data collection process. The

survey is being sent during the final quarter of the school year. This may introduce response

bias as teachers may be fatigued at the end of a given school year and those who respond may

be meaningfully different than those who do not.

The third limitation is related to the self-reported data of the participants. In a review

of the research on the limitation of self-reported data, researchers find that there is no way to

independently verify participant responses (University of Southern California Libraries

Research Guides, n.d.). Therefore, all responses must be taken at face value.

Summary

In conclusion, this chapter provided an overview of the methodology included for the

study of elementary tenured teachers' perception of standards-based teacher evaluation

experiences. The study replicates Pizzi's (2009) instrument, where he developed and tested

the Teacher Evaluation System Attitude Scale. The study involves a mixed methods approach.

The chapter also described the participants, human subject approval, research design,

instruments used for data collection and analysis, procedures and timeline, and limitations of

the study. The results of the study are included in Chapter IV.

Chapter IV: Findings and Results

Introduction

Federal legislation including No Child Left Behind (NCLB, Public Law 107-110) and the Race to The Top (RTTT, Public Law 111-115) federal grant stimulus have provided impetus for reform and innovation in teacher evaluation throughout the United States, including in the state of Minnesota. This intensified focus on accountability for teachers and principals aimed at increasing student achievement has contributed to many school districts designing effective teacher evaluation systems that promote and encourage a culture of continuous improvement and growth for all teachers. Accompanying the design of effective teacher evaluation systems is the perceived need to question teachers about their understanding of the most essential elements of teacher evaluation that are intended to increase their effectiveness.

Chapter II reviewed the related literature on standards-based teacher evaluation. Three themes emerged from the research: (1) measuring teacher effectiveness, (2) standards-based teacher evaluation models, and (3) teacher conferencing and feedback. The research revealed the importance of developing teaching standards that reflect a comprehensive view of teaching (MET, 2013). Additionally, research on factors that contribute to quality feedback identified descriptive feedback (Feeney, 2007), paired with a performance rubric, and a systematic structure for teacher conferencing (face-to-face) and feedback led to teachers reporting greater confidence in the teacher evaluation process.

Purpose of the Study

The study sought to collect data on the perceptions of tenured elementary school teachers in a Minnesota metropolitan school district on the accuracy and usefulness of

standards-based teacher evaluation on teacher effectiveness. Further, the study sought to identify the important components of conferencing (face-to-face) and written feedback within the teacher evaluation process that lead to instructional improvement. The study was conducted in May, 2017. Chapter I included an introduction of the research problem and the purpose of the study. Chapter II provided a review of the related literature of standards-based teacher evaluation organized into three themes. Chapter III included a description of the methodology used to collect data on elementary tenured teachers' standards-based teacher evaluation experiences. Chapter IV reports the results of the study.

Survey and Participants

The researcher received approval from the doctoral committee and the Internal Review Board of St. Cloud State University. In addition, the study received approval from the metropolitan school district's Department of Research, Evaluation and Assessment. The researcher collaborated with the Department of Human Resources to identify tenured elementary school teachers from randomly selected schools. The Department of Research, Evaluation and Assessment encouraged the researcher to extend the number of schools to thirteen to increase the potential number of responses collected. The researcher accepted the recommendation. As a result, there were a potential of 340 participants.

On May 18, 2017, an email (see Appendix E) accompanying a phone call to each of the principals from the thirteen elementary schools was completed to encourage a favorable response rate. An email (see Appendix F) was distributed on May 19, 2017 to potential study participants with a link to the electronic survey. One week later, a reminder email was provided on May 26, 2017. After receiving an email from a potential study participant regarding a problem with the link to the survey, the researcher collaborated with the

metropolitan school district's Department of Research, Evaluation and Assessment and the Statistical Consulting and Research Center from St. Cloud State University to problem solve the issue of study participants not being able to view a link embedded in the original or reminder email. The solution was to create an URL that could be embedded in the email as a link. A final email (see Appendix G) was provided to potential study participants with the correction and the response rate of survey completion increased from 13 to 39. The final response rate on completed surveys was 11.47%. The chapter presents the results of the survey using data tables that provide descriptive and correlational analyses. The initial data presentation analyzes the demographic information related to survey participants. The survey findings were sequenced to match each of the four research questions.

Research Questions

1. How did a select sample of tenured elementary school teachers from a Minnesota metropolitan school district differ in their perceptions of the accuracy of standards-based teacher evaluation in measuring teacher effectiveness?

2. What did a select sample of tenured elementary school teachers in a Minnesota metropolitan school district identify as useful components of a standards-based teacher evaluation system?

3. What did a select sample of tenured elementary school teachers in a Minnesota metropolitan school district identify as those outcome(s) of their standards-based teacher evaluation (including teacher conferencing and written feedback) that were most beneficial to them in improving their teacher effectiveness?

4. How did a select sample of tenured elementary school teachers in a Minnesota metropolitan school district perceive written feedback from standards-based teacher evaluation of value in impacting their professional growth?

Survey Results: Participant Demographics

Demographic information was collected for each of the survey respondents. Information collected did not include items that would compromise the anonymity of the respondents.

Table 4.1 presents information about survey respondents' years of experience as a teacher. Survey responses indicated that 13 respondents or 40.1 % had 21 or more years of experience (Table 4.1). Only 8 respondents or 25.0 % reported that they had been teaching from 16 to 20 years.

Table 4.1

Reported Years of Experience as Teacher

Years of Experience Range	n	Percent
4-10	6	18.8
11-15	5	15.6
16-20	8	25.0
21+	13	40.6
Total	32	100.0

Table 4.2 identifies information collected on survey respondents' field of instruction. Responses indicate that 25 respondents or 78.1% served as K-5 classroom teachers (Table 4.2). Six respondents or 18.8% identified Special Education as their field of instruction.

Table 4.2

Reported Field of Instruction

Field of Instruction	n	Percent
K-5	25	78.1
Special Education	6	18.8
English as Second Language	1	3.1
Total	32	100

In addition to identifying years of experience and field of instruction, survey participants were asked to identify their race. As reported in Table 4.3, 29 respondents or 74.4% identified White as their race. Only one teacher or 3.1% identified himself/herself as Hispanic or Latino, Asian, and Native American or Alaskan Native.

Table 4.3

Respondents' Reported Race

Race	n	Percent
White	29	90.6
Hispanic or Latino	1	3.1
Asian	1	3.1
Native American or Alaskan Native	1	3.1
Total	32	100.0

Survey Results: Research Question One

How did a select sample of tenured elementary school teachers from a Minnesota metropolitan school district differ in their perceptions of the accuracy of standards-based teacher evaluation in measuring teacher effectiveness? The first research question sought to

collect data on how teachers' perceive the accuracy of the standards-based evaluation tool

including the feedback provided within the observational cycle. There were two items from

the survey conducted that provided data pertinent to the research question. Item six requested

survey participants to rate the use of the performance rubric (and its descriptors) as the

common language for effective teaching at their schools. Item eight invited respondents to

rate administrators' knowledge and understanding of effective teaching.

Tables 4.4 and 4.5 reveal the frequency results for items six and eight. The frequency

results from item six as reported in Table 4.4 identified that 21 survey participants or 60.0%

agreed or strongly agreed that the descriptors for the standards of effective teaching serve as

the common language for effective teaching at their schools. The frequency results from item

eight as reported in Table 4.5 identify that 23 survey participants or 67.7% agreed or strongly

agreed that administrators have substantial knowledge and understanding of effective

teaching.

Table 4.4

Frequency Results for Descriptors for the Standards of Effective Teaching Serve as the

Common Language for Effective Teaching

	Frequency	Valid Percent	Cumulative Percent
SD	1	2.9	2.9
D	13	37.1	40.0
A	16	45.7	85.7
SA	5	14.3	100.0
Total	35	100.0	
Missing	4		
Total	39		

Table 4.5

Frequency Results for Administrators Have Substantial Knowledge and Understanding of

Effective Teaching

	Frequency	Valid Percent	Cumulative Percent
SD	1	2.9	2.9
D	10	29.4	32.4
A	12	35.3	67.6
SA	11	32.4	100
Total	34	100.0	
Missing	5		
Total	39		

Tables 4.6 and 4.7 report the mean results by years of experience for items six and

eight of the survey. Table 4.6 identified a mean score of 2.6875 for survey participants'

response to the descriptors for the standards of effective teaching serve as the common

language for effective teaching at their schools. Table 4.7 revealed a mean score of 3.0000 for

survey participants' response to administrators have substantial knowledge and understanding

of effective teaching. The P-value equaled .705 indicated that there is no statistically

significant differences found and failed to reject the Null Hypothesis.

Table 4.6

Mean Results for Standards of Effective Teaching

Years of Experience	n	Mean	Std. Deviation
4-10	6	3.0000	.89443
11-15	5	2.6000	.89433
16-20	8	2.5000	.53452
21+	13	2.6923	.85485
Total	32	2.6875	.78030

Note: P-value = .705. No significant differences were found; fail to reject the Null Hypothesis.

Table 4.7

Mean Results for Administrators Have Substantial Knowledge and Understanding of Effective

Teaching

Years of Experience	n	Mean	Std. Deviation
4-10	6	3.3333	.81650
11-15	5	3.0000	1.0000
16-20	8	2.6250	.91613
21+	13	3.0769	.86232
Total	32	3.0000	.87988

Note: P-value = .511. No significant differences were found; fail to reject the Null Hypothesis.

Table 4.8

Independent Samples T-test, Survey Items Six and Eight

Item	t	Significance
The Standards of Effective Teaching and their descriptors serve as the common language of effective teaching in my school.	.071	.126
Administrators in my school communicate a substantial understanding of effective teaching.	.509	.429

Table 4.8 presents data from an Independent Samples t-test to compare the responses

of teachers who identified their field of instruction as K-5 to those who identified as other

(special education or English as a second language). Based on the P-values shown on Table

4.8, there was not a statistically significant difference in the response by field of instruction

for items six and eight. The P-Value for item six and eight were, respectfully, .126 and .429.

Since both values were greater than .05, there are no findings of statistically significant

differences in the manner in which teachers responded to items six and eight based on their

field of instruction.

Table 4.9

ANOVA, Survey Items Six and Eight

Item	F	Sig.
The standards of effective teaching and their descriptors serve as the common language of effective teaching in my school.	.470	.705
Administrators in my school communicate a substantia understanding of effective teaching.	.788	.511

Table 4.9 presents the ANOVA test employed to compare the responses of survey

participants by years of experience to identify if there were significant statistical differences

in the manner in which the four groups responded to items six and eight. Based on the

findings that both items had P-values greater than .05, item six had a P-value of .705 and item

eight had a P-value of .511, it was determined there were no statistically significant

differences in the responses of each of the four groups.

Survey Results: Research Question Two

What did a select sample of tenured elementary school teachers in a Minnesota

metropolitan school district identify as useful components of a standards-based teacher

evaluation system? Research question two sought to collect information from survey

participants on the components of teacher evaluation that were useful to them. In the second

section of the survey, respondents rated the importance of evaluation activities including

conferencing and written feedback. Two survey items, nine and ten aligned with research

question two. Table 4.10 and Table 4.11 revealed the frequency results for teachers' responses

to items nine and ten. Thirty-four or 100.0% of the survey respondents who completed item

nine identified that conferencing (face-to-face) was an important component of teacher

evaluation. Thirty-three or 97.1% of the survey respondents identified written feedback was

an important component of teacher evaluation.

Table 4.10

Frequency Results for Conferencing is an Important Component of Teacher Evaluation

	Frequency	Valid Percent	Cumulative Percent
A	21	61.8	61.8
SA	13	38.3	100.0
Total	34	100.0	
Missing	5		
Total	39		

Table 4.11

Frequency Results for Written Feedback is an Important Component of Teacher Evaluation

	Frequency	Valid Percent	Cumulative Percent
D	1	2.9	2.9
A	22	64.7	67.6
SA	11	32.4	100.0
Total	34	100.0	
Missing	5		
Total	39		

Tables 4.12 and 4.13 identify the mean results for survey respondents on items nine

and ten based on years of experience. When analyzing the ANOVA test results to compare the

four respondent groups based on their years of experience, the P-values for item nine and item

ten, respectfully, were .640 and .394. Both values were found to be greater than .05. Based on

these findings, there were no statistically significant differences in the responses to items nine

and ten related to conferencing and written feedback based on years of teaching experience.

Table 4.12

Mean Results for Conferencing (face-to-face) between Teachers and Administrators is an

Important Component of Teacher Evaluation

Years of Experience	n	Mean	Std. Deviation
4-10	6	3.5000	.54772
11-15	5	3.2000	.44721
16-20	8	3.5000	.53452
21+	13	3.3077	.48038
Total	32	3.3750	.49187

Note: P-value = .640 and is greater than .05; failed to reject the Null Hypothesis.

Table 4.13

Mean Results for Written Feedback Provided by Administrators is an Important Component

of Teacher Evaluation

Years of Experience	n	Mean	Std. Deviation
4-10	6	3.5000	.54772
11-15	5	3.2000	.44721
16-20	8	3.5000	.53452
21+	13	3.1538	.55470
Total	32	3.3125	.53506

Note: P-value = .394 and is greater than .05; failed to reject the Null Hypothesis.

Table 4.14 presents the data collected from the Independent Sample t-tests for survey items nine and ten. The independent samples t-test compared the survey participant responses of teachers who identified K-5 as their field of instruction compared to other fields of instruction (Special Education and English as a Second Language). There was a statistically significant difference found in the responses of survey participants in their rating of item nine, based on the P-value of .036 which is less than .05. This was determined with a 95% confidence interval with a lower end of .03227 and an upper end of .89081. With a P-Value of

.114 a value greater than .05, there was no statistically significant difference in the way survey participants responded to item ten based on field of instruction.

Table 4.14

Independent Samples T-test for Survey Items Nine and Ten

Item	t	Significance
Conferencing (face-to-face) between teachers and administrators is an important component of teacher evaluation.	2.196	.036
Written feedback provided by administrators is an important component of teacher evaluation.	1.629	.114

Survey Results: Research Question Three

What did a select sample of tenured elementary school teachers in a Minnesota metropolitan school district identify as those outcome(s) of their standards-based teacher evaluation (including teacher conferencing and written feedback) that were most beneficial to them in improving their teacher effectiveness? Data collected from section three of the survey provided information on the extent to which survey respondents affirmed that teacher evaluation influenced instructional improvement. Survey items 15, 17 and 18 all provided information aligned to research question three. Table 4.15, Table 4.16 and Table 4.17 reported the frequency results for items 15, 17 and 18. Nineteen or 57.6% of respondents identified they strongly disagreed or disagreed that the process of teacher evaluation led to increased student achievement at their school. Twenty-one or 63.7% of respondents strongly agreed or agreed that written feedback helped improve the quality of their teaching. Thirty or 90.9% of respondents strongly agreed or agreed that when completed, conferencing (face-to-face) between teachers and administrators as a part of teacher evaluation was helpful in improving teaching and learning.

Table 4.15

Frequency Results for the Process of Teacher Evaluation Leads to Increased Student

Achievement at My School

	Frequency	Valid Percent	Cumulative Percent
SD	3	9.1	9.1
D	16	48.5	57.6
A	11	33.3	90.9
SA	2	6.1	97.0
NA	1	3.0	100.0
Total	33	100.0	
Missing	6		
Total	39		

Table 4.16

Frequency Results for the Written Feedback I Get From Evaluation Helps Me Improve the

Quality of My Teaching

	Frequency	Valid Percent	Cumulative Percent
SD	2	6.1	6.1
D	9	27.3	33.3
A	12	36.4	69.7
SA	9	27.3	97.0
NA	1	3.0	100.0
Total	33	100.0	
Missing	6		
Total	39		

Table 4.17

Frequency Results for Conferencing Between Teachers and Administrators is Helpful in

Improving Teaching and Learning

	Frequency	Valid Percent	Cumulative Percent
D	1	3.0	3.0
A	16	48.5	51.5
SA	14	42.4	93.9
NA	2	6.1	100.0
Total	33	100.0	
Missing	6		
Total	39		

Table 4.18 presents the results from the Independent Samples t-test conducted for

items 15, 17 and 18. The Independent Sample t-tests compared the responses of survey

participants who identified themselves as K-5 teachers to Other teachers (Special Education

and English as a Second Language). Based on the P-values for each item, there were no

statistically significant differences. The P-values included the following: item 15 equaled

.180, item 17 equaled .535, and item 18 equaled .143. None of these P-values was less than

.05, indicating there is a failure in rejecting the Null Hypothesis.

Table 4.18

Independent Samples T-test, Survey Items Fifteen, Seventeen and Eighteen

Item	t	Significance
In my opinion, the process of teacher evaluation leads to increased student achievement at my school.	1.374	.180
The written feedback I get from evaluation helps me improve the quality of my teaching	-.628	.535
When done well, conferencing (face-to-face) between teachers and administrators as a part of teacher evaluation is helpful in improving teaching and learning.	1.506	.143

The ANOVA test presented in Table 4.12 was completed for items 15, 17 and 18 in which the responses of survey participants were compared by their years of teaching experience. Based on the P-values for each of the three items, there were no statistically significant differences in the responses of survey participants based on their years of teaching experience. The P-value for each included the following: item 15 equaled .415, item 17 equaled .411, and item 18 equaled .676. The ANOVA test results which resulted in P-values of greater than .05 for each of the items tested resulted in a failure to reject the Null Hypothesis.

Table 4.19

ANOVA Test for Items Fifteen, Seventeen and Eighteen

Item	F	Sig.
In my opinion, the process of teacher evaluation leads to increased student achievement at my school.	.982	.415
The written feedback I get from evaluation helps me improve the quality of my teaching.	.991	.411
When done well, conferencing (face-to-face) between teachers and administrators as a part of teacher evaluation is helpful in improving teaching and learning.	.514	.676

Survey Results: Research Question Four

How did a select sample of tenured elementary school teachers in a Minnesota metropolitan school district perceive written feedback from standards-based teacher evaluation of value to them in impacting their professional growth? Research question four sought to collect data from respondents on the impact of written feedback on their professional growth. There were two items from the survey that provided information for

research question four. Item 17 and the first open ended response question, item 23. Item 17 was also examined in research question three.

As reported previously, 21 or 63.7% of survey participants identified that written feedback helped them improve the quality of their teaching. Based on the data collected from the Independent Sample t-test and the ANOVA presented in Tables 4.18 and 4.19, there were no statistically significant differences in the responses of survey participants based on their field of instruction or years of teaching experience. Item 23 was an open ended response item that sought to collect data from respondents on the most useful components of conferencing and written feedback. Eighteen or 55.8% of respondents identified that conferencing and written feedback which promoted teacher reflection and led to instructional improvements were the most useful within teacher evaluation. Five participants or 15.5% of respondents identified that the most useful component of conferencing and written feedback was the use of the performance rubric descriptors (or the Standards of Effective Teaching).

Table 4.20

Most Useful Components of Conferencing and Written Feedback Identified by Survey Participants

Useful Component	n	Valid Percent
Specific feedback based on observed data	9	27.9
Promoted reflection & leads to instructional improvements	18	55.8
Performance Rubric Descriptors & Other	5	15.5
Total	32	100

A Chi-square test was completed to compare survey participants' open ended responses from item 23 based on their field of instruction and years of teaching experience. The expected counts for the Chi-square test were less than 5 for 50.0% of the cells. As a

result, this test was used with caution. The P-value for comparing the fields of instruction to most useful component was .391 which is greater than .05. Since there was not a statistically significant difference in the responses based on field of instruction, this resulted in a failure to reject the null hypothesis. The P-value for comparing years of teaching experience compared to most useful component was .663, a figure greater than .05. Therefore, there was no statistically significant difference in teachers' responses based on their years of teaching experience compared to most useful component. This resulted in a failure to reject the null hypothesis.

Conclusion

Chapter IV reported the findings and results about the perceptions of tenured elementary school teachers from a Minnesota metropolitan school district regarding their experiences with standards-based teacher evaluation. Specifically, respondents' perceptions were gathered on the usefulness, accuracy, and components of teacher evaluation including conferencing and written feedback that influenced instructional improvement.

The result of the Cronbach's Alpha analysis was a .065 alpha coefficient with a 95% Confidence Interval. According to the Statistical Research and Consulting Center at Saint Cloud State University, this was an acceptable alpha coefficient for meeting the standards for valid and reliable data.

Survey participants responded to 22 Likert Scale and two open-response items. Seven of the Likert Scale items provided data for the four research questions posed in the study. One of the two open ended response questions provided information for research question four. Based on the Independent Samples t-Test conducted for all Likert Scale items of the survey, only item nine was found to have a statistically significant difference in the manner in which

teachers responded based on their field of instruction. As reported earlier in the chapter, the P-value for item nine was .036 which is less than .05 making it a statistically significant difference. This was determined with a 95% Confidence Interval with the difference on the lower end being .03227 and the upper end being .89081.

Of the 22 Likert Scale items asked of survey participants, item one, which identified respondents' familiarity with the standards of effective teaching, had the highest mean average of 3.8286 with a standard deviation of .45282. Item fifteen, which indicated respondents' level of agreement that the process of teacher evaluation leads to increased student achievement, had the lowest mean average of 2.4545 with a standard deviation of .86930. Items two and three, which identified respondents' understanding of the role of the standards of effective teaching as well as eighteen and twenty-two, which identified the role that conferencing and reflection had on instructional improvement, all had mean averages of greater than or equal to 3.5000.

The frequency results provided several important findings. Twenty-one or 60.0% of respondents strongly agreed or agreed that the performance rubric served as the common language for effective teaching at their school. Thirty-four or 100.0% of respondents strongly agreed or agreed that conferencing (face-to-face) was an important component of teacher evaluation. Thirty-three or 97.1% of respondents strongly agreed or agreed that written feedback was an important component of teacher evaluation. Nineteen or 57.6% of respondents strongly disagreed or disagreed with the statement that the process of teacher evaluation leads to improved student achievement at their school. Thirty or 90.9% of respondents strongly agreed or agreed that when completed, conferencing (face-to-face) was helpful in improving teaching and learning.

The Chi-square test completed for item twenty-three revealed no statistically significant difference in the ways that teachers identified the most useful components of conferencing and written feedback. Eighteen teachers or 55.8% of the survey participants reported that experiences within conferencing (face-to-face) were the most useful within teacher evaluation.

Chapter V will further examine the findings and results from Chapter IV, provide recommendations for the field, and suggestions for future research.

Chapter V: Summary

Study Overview

The study examined the perceptions of tenured elementary teachers from a Minnesota metropolitan school district on the accuracy and usefulness of standards-based teacher evaluation in measuring teacher effectiveness. Additionally, the study sought to identify how teacher conferencing and written feedback within standards-based teacher evaluation influenced their professional growth (development).

The results of the mixed methods study advanced the research discussion on perceptions of standards-based teacher evaluation. Specifically, the results provide data related to the influential components of teacher conferencing and written feedback within standards-based teacher evaluation. The results also provided insights for the field of practice regarding the critical role that administrators play in facilitating effective teacher evaluation practices that lead to instructional improvement.

Chapter V presents a summary of the study including conclusions from the data findings presented in Chapter IV, limitations, recommendations for future research, and recommendations for professional practice.

Research Questions

The research study included four research questions that influenced adaptations to a mixed methods survey tool originally created in a study by Jonathon Pizzi (2009).

1. How did a select sample of tenured elementary school teachers from a Minnesota metropolitan school district differ in their perceptions of the accuracy of standards-based teacher evaluation in measuring teacher effectiveness?

2. What did a select sample of tenured elementary school teachers in a Minnesota metropolitan school district identify as useful components of a standards-based teacher evaluation system?

3. What did a select sample of tenured elementary school teachers in a Minnesota metropolitan school district identify as those outcome(s) of their standards-based teacher evaluation (including teacher conferencing and written feedback) that were most beneficial to them in improving their teacher effectiveness?

4. How did a select sample of tenured elementary school teachers in a Minnesota metropolitan school district perceive written feedback from standards-based teacher evaluation of value to them in impacting their professional growth?

Research Findings: Question One

How did a select sample of tenured elementary school teachers from a Minnesota metropolitan school district differ in their perceptions of the accuracy of standards-based teacher evaluation in measuring teacher effectiveness? Research conducted by Heneman et al. (2006) studied four school districts' implementation of standards-based teacher evaluation systems. Their study revealed that both administrators and teachers reported "most positive and least varied" when responding to the performance competency model within the evaluation system (Heneman et al., 2006, p. 6). Further they reported, "most teachers perceived the ratings as fair and accurate" (Heneman et al., 2006, p. 6). These findings from the literature are similar to the study's results.

The frequency results from item six revealed that twenty-one or 60.0% of survey participants strongly agreed or agreed that the performance rubric or standards of effective teaching serve as the common language for effective teaching at their school.

Item eight frequency results identified that 23 or 67.3% of survey participants strongly agreed or agreed that administrators in their school communicate a substantial understanding of effecting teaching. Additionally, this is further confirmed when examining the mean averages for item six and eight. The average mean for item six was 2.6875 with a standard deviation of .78030. When analyzing years of experience in response to this item, each of the four groups had a mean average above 2.5000. The average mean for item eight was slightly higher at 3.0000 with a standard deviation of .87988. Three of the four groups for years of experience had an average mean at or above 3.000. These findings suggest that the 32 tenured elementary teachers who completed the survey tended to agree that the standards-based performance rubric of the metropolitan school district serves as the common language of effective teaching in their school. They also tended to agree that administrators in their school communicate a substantial understanding of effective teaching.

Research Findings: Question Two

What did a select sample of tenured elementary school teachers in a Minnesota metropolitan school district identify as useful components of a standards-based teacher evaluation system? Anast-May et al. (2011) studied the perceptions of teachers' experiences with standards-based teacher evaluation in a large South Carolinian school district that included 37 elementary teachers. They report, "all 37 participants responded that nothing takes the place of face-to-face conferencing nor can the same information be relayed in an email or a little note left on the desk after the observation" (Anast-May et al., 2011, p. 5). This finding from the literature was confirmed in the study.

Items nine and ten requested that survey participants identify the extent to which they agreed that conferencing and written feedback were important components of standards-based

teacher evaluation. Thirty-four or 100.0% of the survey participants strongly agreed or agreed that conferencing (face-to-face) was an important component of teacher evaluation. Thirty-three or 97.1% of survey participants strongly agreed or agreed that written feedback was an important component of teacher evaluation. The mean for both of these items exceeded 3.0000. Item nine had a mean average of 3.3750 and item ten had a mean average of 3.3125. The mean averages for items nine and ten suggested that teachers agreed with the statements that conferencing and written feedback were important components of standards-based teacher evaluation.

When comparing survey participants responses to item nine by field of instruction, there was a statistically significant difference in the way that K-5 teachers responded to conferencing and written feedback than special education and English as a Second Language teachers. The P-value of .036 for item nine was less than .05. This was determined with a 95% confidence interval with a lower end of .03227 and an upper end of .89081. The literature reviewed did not provide any prior statistical analysis related to the comparison of K-5 elementary classroom teachers to special education and English as a Second Language teachers on conferencing and written feedback.

Research Findings: Question Three

What did a select sample of tenured elementary school teachers in a Minnesota metropolitan school district identify as those outcome(s) of their standards-based teacher evaluation (including teacher conferencing and written feedback) that were most beneficial to them in improving their teacher effectiveness? The literature on the influence of standards-based teacher evaluation on instructional improvement affirmed the importance of essential criteria related to constructive feedback. Feeney (2007) proposed that constructive feedback

should include the following criteria: 1) be based on descriptive observable data (Danielson & McGreal, 2000); 2) provide characteristics of effective teaching (Danielson, 1996; Marzano et al., 2001); and 3) promote reflective inquiry and self-directedness to foster improvements in teaching supported by evidence of student learning (Glickman, 2002). Feeney (2007) found that the third criteria was hard to measure. Anast-May et al. (2011) reported in their study of elementary teachers' perceptions of standards-based teacher evaluation experiences that "face-to-face discussion is more powerful and has a greater impact" (Anast-May et al., 2011, p. 5). This key finding from the literature regarding the importance of face-to-face conferencing was confirmed in the study.

Thirty or 90.0% of survey participants strongly agreed or agreed that when completed, conferencing (face-to-face) is helpful in improving teaching and learning. When comparing the responses of teachers from item eighteen which asked survey participants to identify if they agreed that face-to-face conferencing was helpful in improving teaching and learning, the mean average for twenty-six K-5 teachers was 3.6154 with a standard deviation of .90469. This was the highest mean average of any of the items for K-5 teachers related to components of teacher evaluation that contributed to instructional improvements. The overall mean average for item eighteen was 3.5313 with a standard deviation of .67127. Based on the data collected from the study, K-5 teachers tended to strongly agree or agree that face-to-face conferencing was helpful in improving teaching and learning.

Research Findings: Question Four

How did a select sample of tenured elementary school teachers in a Minnesota metropolitan school district perceive written feedback from standards-based teacher evaluation of value in impacting their professional growth? Zimmerman and Deckert-Pelton

(2003) completed a study of 86 educators from five northwest Florida counties. They found

that teachers' experiences with written feedback was inconsistent. They further identified that

the evaluator's pedagogical knowledge affected the quality of feedback that teachers received.

Zimmerman and Deckert-Pelton (2003) reported, "the relationship between principal and

faculty members has a pivotal effect on instructional effectiveness" (Zimmerman & Deckert-

Pelton, 2003, p. 29). Valentine (1992) as cited in Zimmerman and Deckert-Pelton (2003)

maintained "that implementing the improvement component is one of the most challenging

tasks of the principal in the performance-based developmental evaluation process" (p. 29).

The literature on written feedback as a component to improving instructional effectiveness

was confirmed in the study.

  The results of item seventeen that asked survey participants to identify the extent to

which they agreed that written feedback helped them improve the quality of their teaching

suggested that there may be disparate experiences of teachers with respect to written feedback

that they receive. Twenty-one or 63.7% of survey participants strongly agreed or agreed that

written feedback they received from evaluation helped improve the quality of their teaching.

Further, the mean average for item seventeen was 2.9375 with a standard deviation of .98169.

This standard deviation suggested that the range of responses from the mean was nearly

1.00000, revealing that experiences with written feedback may have been inconsistent from

teacher to teacher or from school to school.

  When examining the results of item 23, an open ended response item that requested

survey participants to identify the most useful component of conferencing and written

feedback, 55.8 % or 18 teachers identified the following experiences related to conferencing

(face-to-face) as most useful in teacher evaluation: promoted reflection and led to

instructional improvements. This finding confirmed the research previously cited from Anast-May et al. (2011) regarding the power and impact of face-to-face discussion within teacher evaluation.

Summary of Findings

The statistical data analysis conducted for the survey items related to all four research questions found one item with a statistically significant difference in the way participants responded. Item six had a P-value of .036 which was less than .05 resulting in a failure to reject the Null Hypothesis. All other items were not found to have a statistically significant difference.

Limitations

The following are limitations of the study.

1. The sample size or number of survey participants who completed the survey totaled 32 of a possible 340. This small completion rate made it difficult to establish significant relationships within the statistical data analysis.

2. The instrument used to gather data involved replicating an original study from Pizzi (2009). Based on the data findings, the survey items were not as closely aligned to the research questions as they might have been. Additionally, the qualitative portion of the instrument needs to be revised to elicit greater quantities of information that could potentially reveal significant relationships among survey participants.

3. The timing of the distribution of the survey was found to be challenging for teachers. Distribution occurred during the final four weeks of the school year, a time of the school year when teachers were likely fatigued and not disposed to

complete surveys regardless of incentives provided. In addition, the survey posed access issues which caused respondents to be unable to view the link in the email sent from Survey Monkey. The timing of the solution to the access issue further delayed the time available for teachers to complete the survey.

Recommendations for Further Research

The following recommendations for further research address the limitations identified in the previous section.

1. **It is recommended that future research be conducted with an increased sample size,** by including all tenured teachers K-12 in a Minnesota metropolitan school district. Additionally, including multiple Minnesota metropolitan school districts could increase sample size and provide additional data points for further identification of significant relationships.

2. **It is recommended that further research be undertaken to develop a qualitative tool** that uses a focus group to further clarify key components of standards-based teacher evaluation that influence instructional improvements. The revision of the quantitative Likert Scale items should be prioritized to increase alignment with research questions. This would increase the relevance and number of data points to examine in relationship to the research questions posed.

3. **It is recommended that the timing of the study** should be identified in collaboration with the identified school district's research, evaluation and assessment staff members. These district departments have a local sense of the timing that would most likely contribute to a maximum response rate.

Recommendations for Practice

Based on the literature reviewed and the survey findings within the study, it is believed standards-based teacher evaluation has the potential to increase teacher effectiveness. Additionally, teachers are experiencing inconsistent interactions and feedback with the use of standards-based teacher evaluation. School district leaders have the opportunity to implement meaningful teacher evaluation systems that lead to professional learning for all of it's teachers if those leaders apply the lessons that may be gleaned from the related literature or research. The following recommendations provide school district leaders with guidance on how to increase the effective and ongoing implementation of standards-based teacher evaluation.

1. **It is suggested that professional development for all evaluators on constructive feedback should be implemented and required annually**. If conferencing and written feedback are reported by teachers as the most useful components of standards-based teacher evaluation that result in instructional improvement, then evaluators would be advised to acquire the demonstrated capacity to meet the three criteria proposed by Feeney (2007) for constructive feedback. In addition, pedagogically appropriate feedback must be integrated into evaluator's professional development to ensure the relevance and applicability of the feedback that will be provided to teachers across all licensure areas.

2. **It is suggested that inter-rater reliability must be a key component in maintaining the fidelity of standards-based teacher evaluation**. Inter-rater reliability needs to be a district-led commitment to ensure that teachers have evaluation experiences that are consistent and fair across all schools in the school district.

3. **It is suggested that school district leaders increase the frequency of observations of tenured teachers to greater than one observation each school year**. Research by Anast-May et al. (2011) found that when teachers trusted the evaluation feedback and regarded the evaluator as coach they would welcome being observed at least once each quarter of the school year. The frequency of quality feedback to teachers must be more frequent than once a school year to support the ongoing professional learning of teachers.

Summary

The study examined the perceptions of tenured elementary school teachers from a Minnesota metropolitan school district on the accuracy of standards-based teacher evaluation in measuring teacher effectiveness. The study sought to further examine the perceptions of respondents related to the useful components of conferencing and written feedback within the teacher evaluation process. Based on the findings from the literature and the data collected in the study, it is clear that standards-based teacher evaluation has the potential to accurately measure teacher effectiveness and contribute to instructional improvements for teachers. These potential outcomes align to the purpose of teacher evaluation as described by Maslow and Kelley (2012), they reported that teacher evaluation should provide "meaningful feedback to teachers to improve teacher practice" and to be an "important source of data to inform organizational systems that support teaching and learning" (p. 601).

The study findings identified that teachers agreed that conferencing is the most useful component of standards-based teacher evaluation. They also tended to agree that the performance rubric descriptors provided a common language for teachers and administrators in their schools. As a result, it would appear to be imperative that evaluators need to possess

the pedagogically appropriate understanding of curriculum and instruction that enables them to assess high quality teaching across academic disciplines.

However, as the literature has maintained, teacher evaluation has not consistently provided quality feedback to teachers (Anast-May et al., 2011; Feeney, 2007, Frase, 2001; Zimmerman & Deckert-Pelton, 2003). Maslow and Kelley (2012) also found this to be true when they reported, "that the quality of evaluation suffers from school leaders' inconsistent implementation, competing demands, and lack of clear understanding of how to assess high-quality teaching" (p. 600).

The findings of the study also confirmed the challenges of teacher evaluation identified in the literature. Based on the standard deviations across the survey items examined in relationship to the research questions, teachers may be experiencing inconsistent interactions with standards-based teacher evaluation from school to school and within the same school. As the study found, 29 or 87.9% of survey participants strongly agreed or agreed that administrators need continuous professional development on teacher evaluation to improve their practice of conducting quality evaluation of teachers.

In conclusion, the study results contribute to advancing the field of research related to teacher perceptions of the value of standards-based teacher evaluation. Conferencing (face-to-face) as identified by teachers is the most useful component of teacher evaluation. This important study finding will hopefully encourage further development of conferencing structures by school districts—ensuring that teacher evaluation further realizes its fullest potential.

References

Anast-May, L., Penick, D., Schroyer, R., & Howell, A. (2011). Teacher conferencing and

feedback: Necessary but missing! *International Journal of Educational Leadership*

Preparation, 6(2), 1-7.

Bill & Melinda Gates Foundation. (2013). *Ensuring fair and reliable measures of effective*

teaching, culminating findings from the measures of effective teaching project. Seattle,

WA: Author.

Callahan, K., & Sadeghi, L. (2015). Teacher perceptions of the value of teacher evaluations:

New Jersey's AVHIEVE NJ. *NCPEA International Journal of Educational*

Leadership Preparation, 10(21), 46-59.

Campbell, R. J., Kyriakides, L., Muijs, D., & Robinson, W. (2004). *Assessing teacher*

effectiveness: Developing a differentiated mode. London: Routledge Falmer.

Danielson, C. (1996). *Enhancing professional practice, a framework for teaching*.

Alexandria, VA: Association for Supervision and Curriculum Development.

Danielson, C. (2007). *Enhancing professional practice: A framework for teaching* (2nd ed.).

Alexandria, VA: Association for Supervision and Curriculum Development.

Danielson, C. (2011). *Enhancing professional practice: A framework for teaching* (Rev. ed.).

Alexandria, VA: Association for Supervision and Curriculum Development.

Danielson, C. (2013). *The framework for teaching evaluation instrument*. Princeton, NJ: The

Danielson Group.

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

Alexandria, VA: Association for Supervision and Curriculum Development.

Donaldson, M. J. (2009, June). *So long, Lake Wobegon? Using teacher evaluation to raise teacher quality*. Washington, DC: Center for American Progress.

Feeney, E. J. (2007). Quality feedback: The essential ingredient for teacher success. *Clearing House: A Journal of Educational Strategies, Issues, and Ideas, 80*(4), 191-197.

Frase, L. E. (2001). Constructive feedback on teaching is missing. *Education, 113*(2), 176-180.

Glickman, C. D. (2002). *Leadership to learning: How to help teachers succeed*. Alexandria, VA: Association for Supervision and Curriculum Development.

Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality, U.S. Department of Education. Retrieved July 9, 2016, from http://dev-tqsource.airws.org/publications/EvaluatingTeachEffectiveness.pdf.

Heneman III, H. G., Milanowski, A., Kimball, S., & Odden, A. (2006). *Standards-based teacher evaluation as a foundation for knowledge- and skill-based pay* (CPRE Policy Briefs No. RB-45). Philadelphia: Consortium for Policy Research in Education. Retrieved July 9, 2016, from http://repository.upenn.edu/cgi/viewcontent.cgi?article= 1017&context=cpre_policybriefs.

Kahlenberg, R. D. (2015). Tenure: How due process protects teachers and students. *American Educator, 39*(2), 4-43.

Kleinman, G. S. (1966). Assessing teaching effectiveness: The state of the art. *Science Education, 50*(3), 234-238.

Leithwood, K., & Louis, K. S. (2011). *Linking leadership to student learning*. San Francisco, CA: Josey-Bass.

Marzano, R. J. (2012). Teacher evaluation: What's fair? What's effective? *Educational Leadership, 70*(3), 14-19.

Marzano, R. J., Pickering, D., & Pollock, J. E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Alexandria, VA: Association for Supervision and Curriculum Development.

Maslow, V. J., & Kelley, C. J. (2012). Does evaluation advance teaching practice? The effects of performance evaluation on teaching quality and system change in large diverse high schools. *Journal of School Leadership, 22*, 600-632.

Mead, S., Rotherham, A., & Brown, R. (2012). *The hangover, thinking about the unintended Consequences of the nation's teacher evaluation binge*. Special Report 2. Washington, DC: American Enterprise Institute.

Minnesota Department of Education. (2017). *Districts, schools, and educators section*. Retrieved from www.education.state.mn.us/**MDE**/SchSup/ESEAFlex/**MMR**/ index.html.

Minnesota state teacher development, evaluation and peer support model evaluation report. (2015, January). Minneapolis: MN: Center for Applied Research and Educational Improvement, College of Education and Human Development, University of Minnesota.

Minnesota Statute: 122A.41 Teacher Tenure Act; Cities of the First Class; Definitions. (2015). State of Minnesota, Office of the Revisor of Statutes.

National Board for Professional Teaching Standards. (2015). *What teachers should know and be able to do.* Arlington, VA: Author.

National Staff Development Council. (2016). Professional development. Oxford, OH: Author. Retrieved from http://www.aypf.org/documents/62609NSDCDefinitionof ProfessionalDevelopment908.pdf.

Nolan, J. H., & Hoover, L. A. (2008). *Teacher supervision & evaluation: Theory into practice* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Pizzi, J. D. (2008). *Urban secondary teachers' perceptions of standards-based teacher evaluation system*. (unpublished doctoral dissertation). Boston, College, Chestnut Hill, MA.

Sergiovanni, T. J., & Starratt, R. J. (2007). *Supervision: A redefinition*. New York, NY: McGraw-Hill.

Simon, M. K. (2011). *Dissertation and scholarly research: Recipes for success.* Seattle, WA: Dissertation Success LLC.

Toch, T. (2008). Fixing teacher evaluation. *Educational Leadership, 66*(3), 32-37.

Tyler, J. H., Taylor, E. S., Kane, T. J., & Wooten, A. L. (2010). Using student performance data to identify effective classroom practices. *American Economic Review, 100*(2), 256-260.

University of Southern California Libraries Research Guides. (on-line). Available: http://libguides.usc.edu/writingguide/limitations.

U.S. Department of Education. (2007). Final guidance on maintaining, collecting, and reporting racial and ethnic data to the U.S. Department of Education. *Federal Register*, 72(202), 59266.

Waters, J. T., Marzano, R. J., & McNulty, B. (2003). *Balanced leadership: What 30 years of research tells us about the effect of leadership on student achievement*. Aurora, CO: Mid-continent Research for Education and Learning.

Zimmerman, S., & Deckert-Pelton, M. (2003). Evaluating the evaluators: Teachers' perceptions of the principal's role in professional evaluation. *NASSP Bulletin, 87*(636), 28-37.

Appendix A: Research Approval

May 11, 2017

Bryan Bass 810 Palace Ave Saint Paul, MN 55439

RE: The Value of Standards-Based Teacher Evaluation: Perceptions of Elementary Tenured Teachers

Dear Mr. Bass:

The Research Committee is pleased to inform you that your research proposal has been reviewed and approved. The approved dates for your research are May 15, 2017 to June 1, 2017. Please submit a copy of your final report to this office within 90 days of completion of the research.

If you have any questions or concerns, please contact me at 651-767-8116. I wish you all the best in your studies.

Sincerely,

Marian Heinrichs, Ph.D Research Committee Co-chair Manager of Program Evaluation Department of Research, Evaluation and Assessment

Appendix B: Pilot Instrument

Teacher Evaluation System Attitude Scale (adapted from Pizzi, 2009)

Directions for Completing Sections B, C & D
Respond to each item in these sections by placing an "X" in the box that best reflects your level of agreement with the item. Please fill in only one box per item. Refer to the chart below when responding.

SD	Strongly Disagree with the item
D	Disagree with the item
A	Agree with the item
SA	Strongly Agree with the item

Section A: Items 1-6, The Standards of Effective Teaching (SET)

SD=**Strongly Disagree** with the item; D=**Disagree** with the item; A=**Agree** with the item; SA=**Strongly Agree** with the item

Item	Description	SD	D	A	SA
1.	**I am familiar with the Standards of Effective Teaching as used in Saint Paul Public Schools.**				
2.	**Our district's teacher evaluation process (pre-conference, observation & post-conference) is wholly based upon the Standards of Effective Teaching.**				
3.	**I understand how the Standards of Effective Teaching document is used in our teacher evaluation process.**				
4.	**In my school, there is regular discussion about what the Standards of Effective Teaching look like in the classroom.**				
5.	**In my school, there is a common understanding among teachers and administrators regarding specific teaching behaviors that model the Standards of Effective Teaching and their descriptors.**				
6.	**The Standards of Effective Teaching and their descriptors serve as the common language of effective teaching in my school.**				

Section B: Items 7-14, Evaluation Activities, Conferencing, and Written Feedback

SD=**Strongly Disagree** with the item; D=**Disagree** with the item; A=**Agree** with the item; SA=**Strongly Agree** with the item

Item	Description	SD	D	A	SA
7.	**In my school, an annual goal setting process for teachers should be linked to teacher evaluation.**				
8.	**Administrators in my school communicate a substantial understanding of effective teaching.**				
9.	**Conferencing (face-to-face) between teachers and administrators is an important component of teacher evaluation.**				

10.	**Written feedback provided by administrators is an important component of teacher evaluation.**				
11.	**In my school, conferences (face-to-face) between teachers and administrators as part of evaluation are done well.**				
12.	**Administrators spend the expected amount of time observing in classrooms as part of the evaluation process.**				
13.	**I believe administrators spend sufficient time in classrooms as part of the evaluation process.**				
14.	**Teachers and administrators in my school share a common understanding of specific criteria for performance ratings ("proficient/distinguished") in the evaluation process.**				

Section C: Items 15-22, Evaluation and Instructional Improvement

SD=**Strongly Disagree** with the item; D=**Disagree** with the item; A=**Agree** with the item; SA=**Strongly Agree** with the item

Item	Description	SD	D	A	SA
15.	**In my opinion, the process of teacher evaluation leads to increased student achievement at my school.**				
16.	**Teacher evaluation is linked closely to my school's professional development initiatives.**				
17.	**The written feedback I get from evaluation helps me improve the quality of my teaching.**				
18.	**When done well, conferencing (face-to-face) between teachers and administrators as a part of teacher evaluation is helpful in improving teaching and learning.**				
19.	**Feedback from peer observation helps me to improve the quality of my teaching.**				
20.	**Administrators need ongoing professional development around teacher evaluation for evaluation to improve my practice.**				
21.	**Multiple sources of student performance data are used as part of my summative evaluation.**				
22.	**I rely upon reflection of my own teaching to improve student performance.**				

Section D: Items 28-31, Open Response Items
Provide a response to the following items Please use only the space provided.

23. Describe the most useful component of the teacher evaluation process at your school.

24. Outline a substantial change you would make to improve the teacher evaluation process at your school.

25. What is a major drawback to using teacher evaluation as a means to improve teaching and learning at your school?

26. Share any additional feedback regarding your experiences with the evaluation process at your school.

Section E: Items 27-32 Professional Demographic Information

Please complete all demographic information below by placing an "X" in the box that best applies. Fill in only one circle per item.

27. Gender: Female   Male

28. Race: Black or African-American   White   Hispanic or Latino or

Spanish Origin   Asian   Native Hawaiian or other Pacific Islander   Native

American or Alaskan Native

Other_____

29. Years of Service in this district: 4-7   8-10   10-20   20+years

30. Years of Service in this building: 1-5   6-10   10-20   20+years

31. Last level of education completed: BA/BS   MA/MS   Doctorate   Post

Doctorate

32. Primary field of instruction: Reading   Writing   Science   Social Studies

Special Education   Physical Education   Arts   ESL/EL, World Language

Other Area_____

Appendix C: Teacher Evaluation System Attitude Scale (adapted from Pizzi, 2009)

<u>Directions for Completing Sections B, C & D</u>
Respond to each item in these sections by placing an "X" in the box that best reflects your level of agreement with the item. Please fill in only one box per item. Refer to the chart below when responding.

SD	Strongly Disagree with the item
D	Disagree with the item
A	Agree with the item
SA	Strongly Agree with the item

<u>Section A: Items 1-6, The Standards of Effective Teaching (SET)</u>

SD=**Strongly Disagree** with the item; D=**Disagree** with the item; A=**Agree** with the item; SA=**Strongly Agree** with the item

Item	Description	SD	D	A	SA
1.	**I am familiar with the Standards of Effective Teaching as used in Saint Paul Public Schools.**				
2.	**Our district's teacher evaluation process (pre-conference, observation & post-conference) is wholly based upon the Standards of Effective Teaching.**				
3.	**I understand how the Standards of Effective Teaching document is used in our teacher evaluation process.**				
4.	**In my school, there is regular discussion about what the Standards of Effective Teaching look like in the classroom.**				
5.	**In my school, there is a common understanding among teachers and administrators regarding specific teaching behaviors that model the Standards of Effective Teaching and their descriptors.**				
6.	**The Standards of Effective Teaching and their descriptors serve as the common language of effective teaching in my school.**				

<u>Section B: Items 7-14, Evaluation Activities, Conferencing, and Written Feedback</u>

SD=**Strongly Disagree** with the item; D=**Disagree** with the item; A=**Agree** with the item; SA=**Strongly Agree** with the item

Item	Description	SD	D	A	SA
7.	**In my school, an annual goal setting process for teachers should be linked to teacher evaluation.**				
8.	**Administrators in my school communicate a substantial understanding of effective teaching.**				
9.	**Conferencing (face-to-face) between teachers and administrators is an important component of teacher evaluation.**				
10.	**Written feedback provided by administrators is an important component of teacher evaluation.**				
11.	**In my school, conferences (face-to-face) between teachers and administrators as part of teacher evaluation are done well.**				

12.	Administrators spend the expected amount of time observing in classrooms as part of the evaluation process.				
13.	I believe administrators spend sufficient time in classrooms as part of the evaluation process.				
14.	Teachers and administrators in my school share a common understanding of specific criteria for performance ratings ("proficient/distinguished") in the evaluation process.				

Section C: Items 15-22, Evaluation and Instructional Improvement

SD=Strongly Disagree with the item; **D=Disagree** with the item; **A=Agree** with the item; **SA=Strongly Agree** with the item

Item	Description	SD	D	A	SA
15.	In my opinion, the process of teacher evaluation leads to increased student achievement at my school.				
16.	Teacher evaluation is linked closely to my school's professional development initiatives.				
17.	The written feedback I get from evaluation helps me improve the quality of my teaching.				
18.	When done well, conferencing (face-to-face) between teachers and administrators as a part of teacher evaluation is helpful in improving teaching and learning.				
19.	Feedback from peer observation helps me to improve the quality of my teaching.				
20.	Administrators need ongoing professional development around teacher evaluation for evaluation to improve my practice.				
21.	Multiple sources of student performance data are used as part of my summative evaluation.				
22.	I rely upon reflection of my own teaching to improve student performance.				

Section D: Items 23-24, Open Response Items
Provide a response to the following items Please use only the space provided.

23. Describe the most useful component of teacher conferencing (face-to-face) and feedback within teacher evaluation.

24. Describe a significant change you would make to improve teacher conferencing (face-to-face) and feedback within teacher evaluation.

Section E: Items 25-27 Professional Demographic Information

Please complete all demographic information below by placing an "X" in the box that best applies. Fill in only one circle per item.

25. Race: Black or African-American   White   Hispanic or Latino or

Spanish Origin   Asian   Native Hawaiian or other Pacific Islander

Native American or Alaskan Native

Other_____

33. Years of Service in teaching: 4-7   8-10   10-20   20+years

34. Primary field of instruction: Reading   Writing   Science   Social

Studies   Special Education   Physical Education   Arts   ESL/EL, World

Language Other Area_____

Appendix D: Invite to Principals

Good morning, (principal):

Thank you for agreeing to support my doctoral study entitled "The Value of Standards-Based Teacher Evaluation: Perceptions'of Elementary Tenured Teachers".

As discussed in our previous conversation, I am requesting you to send the script or language below in an email to the tenured licensed teachers at your school--in support of increasing the participation rate for the study. I have attached a draft of the survey for your convenience.

Greetings, (insert name) Elementary School Teachers:

It is with great enthusiasm that I encourage you to participate in the doctoral study entitled "The Value of Standards-Based Teacher Evaluation: Perceptions'of Elementary Tenured Teachers".

The purpose of the study is to understand the extent to which elementary teachers perceive standards-based teacher evaluation to be accurate and useful in measuring teacher's effectiveness. Additionally, the study seeks to identify how teacher conferencing and written feedback within the standards-based teacher evaluation influences teacher's professional growth. The data from the study will be shared with the Oversight Committee for Teacher Development & Evaluation in support of their commitment to improving the system of teacher evaluation for administrators and teachers.

The survey will be emailed to your district email account this week via Survey Monkey.

Sincerely,

Appendix E: Initial Invite & Consent Statement

The Value of Standards-Based Teacher Evaluation: Perceptions of Tenured Elementary
Teachers in a Metropolitan School District.
Consent to Participate

You are invited to participate in a research study about what tenured elementary teachers
think about standards-based teacher evaluation. Specifically, it explores teachers' thoughts
about whether teacher evaluation methods are effective and contribute to professional growth.
You were chosen as a participant because you are a tenured elementary teacher in a
metropolitan school district that uses standards-based teacher evaluation including the
Standards of Effective Teaching.

If you agree to be part of the research study, you will be asked to complete a survey. The
actual survey will take approximately 20 minutes of your time.

The data will be shared in aggregate with the department of Research, Evaluation and
Assessment and the Oversight Committee for Teacher Development & Evaluation to support
its ongoing improvement of teacher evaluation.

This is not a district sponsored survey. It is also not an evaluation of our evaluation tool or
and evaluation of our principals. There are no risks to participating in the study.

Participating in this study is completely voluntary. Your decision whether or not to participate
will not affect your current or future relations with St. Cloud State University, or the
researcher. If you decide to participate, you are free to withdraw at any time without penalty.

If you have questions about this research study, you may contact Bryan Bass at
bryan.bass@spps.org or Roger Worner at rbworner@stcloudstate.edu. Results of the study
can be requested from the researcher.

If you complete the survey, you will will have the opportunity to enter a raffle for one of three
fifty-dollar gift cards to Target.

Your completion of the survey indicates that you are at least 18 years of age and your consent
to participation in the study.

Appendix F: Final Invite Reminder

Good afternoon,

Sorry for any inconvenience that these additional emails may have caused, I appreciate your willingness to support this research study.

You may have received an email to complete this survey and could not access the link. This email is providing an alternative method to complete the survey. If you have already completed the survey, please ignore this communication.

If you haven't completed the survey and are willing to complete it, please click on the link below. The survey will take approximately seven to ten minutes. There will be three TARGET gift cards of $50 each to be raffled off for all responses submitted.

The Value of Standards-Based Teacher Evaluation Survey

Respectfully,

Bryan E. Bass, St. Cloud State University Doctoral Candidate

Appendix G: IRB Approval

Name: Email:

Bryan Bass bedwardbass@gmail.com

Institutional Review Board (IRB)

720 4th Avenue South AS 210, St. Cloud, MN 56301-4498

IRB PROTOCOL DETERMINATION:

Exempt Review

Project Title: The Value of Standard-Based Teacher Evaluation: Perceptions' of Elementary Tenured Teachers in a Metropolitan School District

Advisor Roger Worner The Institutional Review Board has reviewed your protocol to conduct research involving human subjects. Your

project has been: **APPROVED**

Please note the following important information concerning IRB projects: - The principal investigator assumes the responsibilities for the protection of participants in this project. Any adverse

events must be reported to the IRB as soon as possible (ex. research related injuries, harmful outcomes, significant withdrawal of subject population, etc.).

- For expedited or full board review, the principal investigator must submit a Continuing Review/Final Report form in advance of the expiration date indicated on this letter to report conclusion of the research or request an extension.

-Exempt review only requires the submission of a Continuing Review/Final Report form in advance of the expiration date indicated in this letter if an

extension of time is needed.

- Approved consent forms display the official IRB stamp which documents approval and expiration dates. If a renewal is requested and approved, new consent forms will be officially stamped and reflect the new approval and expiration dates.

- The principal investigator must seek approval for any changes to the study (ex. research design, consent process, survey/interview instruments, funding source, etc.). The IRB reserves the right to review the research at any time.

If we can be of further assistance, feel free to contact the IRB at 320-308-3290 or email ri@stcloudstate.edu and please reference the SCSU IRB number when corresponding.

OFFICE USE ONLY

IRB Institutional Official:

Dr. Latha Ramakrishnan Interim Associate Provost for Research Dean of Graduate Studies

SCSU IRB# 1722 - 2161 1st Year Approval Date: 1st Year Expiration Date:

5/17/2016

Type: Exempt Review **2nd Year Approval Date: 2nd Year Expiration Date:**

Today's Date: 3rd Year Approval Date: 3rd Year Expiration Date: