WALDEN UNIVERSITY
A higher degree. A higher purpose.

Walden University
ScholarWorks

Walden Dissertations and Doctoral Studies

Walden Dissertations and Doctoral Studies Collection

2018

Elementary Principals' Perceptions of Special Education Teachers' Performance Evaluation Process

Nafees H. Rasul
Walden University

Follow this and additional works at: https://scholarworks.waldenu.edu/dissertations

Part of the Special Education Administration Commons, Special Education and Teaching Commons, and the Teacher Education and Professional Development Commons

Walden University

College of Education

This is to certify that the doctoral study by

Nafees Rasul

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Andrea Wilson, Committee Chairperson, Education Faculty
Dr. Kathleen Norris, Committee Member, Education Faculty
Dr. Timothy Lafferty, University Reviewer, Education Faculty

Chief Academic Officer
Eric Riedel, Ph.D.

Walden University
2018

Abstract

Elementary Principals' Perceptions of Special Education Teachers' Performance

Evaluation Process

by

Nafees Rasul

MA, Governors State University, 2003

MA, Governors State University, 2000

MA, Governors State University, 1999

Project Study Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Education

Walden University

April 2018

Abstract

Evidence suggests that current special education teacher evaluation systems may not accurately reflect these teachers' unique duties and responsibilities. In a Midwestern, urban school district, the teacher evaluation system was not adequately aligned with the performance expectations of special education teachers. Guided by Danielson's framework for teaching, this qualitative case study explored elementary school principals' perceptions of teacher evaluation approaches, perceived effectiveness of these approaches, applications of key indicators of teaching quality, and barriers preventing accurate evaluations of special education teachers' performance. An online survey comprised of closed- and open-ended questions was distributed to 445 principals in the district with 97 responding. Descriptive analysis of closed-ended items indicated 70% of respondents perceived current evaluation methods to be insufficiently differentiated for special education teachers' roles and 90% reported a need for additional measures of effectiveness to be used. Thematic analysis of open-ended survey responses confirmed the need for differentiated evaluation approaches to address a misalignment of key effectiveness indicators for special education teachers and revealed barriers to accurate evaluation including resource constraints. In response to these findings, a position paper with policy recommendation prescribed the revision of the current teacher evaluation practices to address the unique roles of special education teachers. Refining special education teacher evaluation practices may contribute to positive social change by aligning the evaluation process with special education teachers' duties and responsibilities, thereby improving teacher performance and potentially increasing student achievement over time.

Elementary Principals' Perceptions of Special Education Teachers' Performance

Evaluation Process

by

Nafees Rasul

MA, Governors State University, 2003

MA, Governors State University, 2000

MA, Governors State University, 1999

Project Study Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Education

Walden University

April 2018

Dedication

To my loving mother and father, whose memory I will always cherish. Your light shone in so many lives. Thank you so much for all your patience and sacrifices and your inspiring example over the years. You taught me the meaning of courage and dedication, and I am grateful to you for encouraging me to follow my dream and pursue a doctoral degree. You always highlighted the importance of higher education and supported my aspirations, and I know that you share this success with me as well. And to the rest of my family, I appreciate all the support and most of all your love.

Table of Contents

iii

vi

List of Tables

Section 1: The Problem

Introduction

In this study I examine how elementary principals perceive the process of evaluating special education teachers' performance. Federal mandates require that state and local school districts evaluate special educators for the role they play in their students' growth, but challenges have been faced in such evaluation (National Comprehensive Center for Teacher Quality [NCCTQ], 2010b). Problems associated with special education teacher evaluation and its description's impact on the local and national population will be considered. This is done through consideration of evidence of the problem, its significance on a local and national level, and the discussion of research questions for this study. The literature review includes various perspectives on current evaluation systems and alternate assessments. I will also present study implications based on the anticipated findings resulting from survey responses with local school district principals.

The Local Problem

Evaluation criteria for special educators must be differentiated from those employed to evaluate general educators (NCCTQ, 2010b). The NCCTQ makes this assertion, noting that "most evaluation systems focus on teacher practice tied to student achievement; however, few systems have the capacity to differentiate among specialty area educators, address the challenges in accurately measuring teacher effects connected to student gains" (NCCTQ, 2010a, p. 1). The principles of teacher evaluation currently show a severe disconnect or gap between the measurement tool and evaluation goals in

schools. The study also identifies a paucity of current research with respect to how special educators are to be evaluated, particularly with respect to tailoring such evaluations to reflect their skills and practice. There is a concern that evaluations are subjective in nature and that the definition of good teaching is unclear (Danielson, 2010). Additional concerns exist in many states and in local schools that evaluation criteria for special educators—which are typically the same as those employed when evaluating general educators—are not applicable for special education teachers who play an expert role instructing special needs students (NCCTQ, 2010). Due to the need to provide individualized support for students with special needs, special educators' instructional practices differ depending on student and school needs (Council for Exceptional Children, 2013). Consequently, evaluations of special educators should reflect this difference.

Rationale

Evidence of the Problem at the Local Level

According to American School and University (2015), school districts have raised concerns about the accuracy of special education teacher evaluations. In the past decade, general teacher evaluation methods across the nation have undergone major changes (Coulter, 2013; Jones & Brownell, 2015). Researchers have shown that this is due to several reasons including ineffective evaluation methods, as is discussed below in the review of the literature (Coulter, 2013; Jones & Brownell, 2015). However, despite these improvements to the general teacher evaluation systems, progress still must be made in determining which evaluation system is most effective for special educators

(Semmelroth, Johnson, & Allred, 2013). The current, updated teacher evaluation methods are geared toward generalized content areas (such as math, reading) and student ability. In other words, no adequate special education teacher evaluation system exists (Semmelroth et al., 2013). Furthermore, researchers proved that current observations have "not sufficiently considered special education," and they implied that there is a need to create effective special teacher evaluation methods (Semmelroth et al., 2013).

Current teacher evaluation systems fail to evaluate the needs of special educators or their students (Lawson & Knollman, 2017; Semmelroth et al., 2013). There are various unmet needs of the special education teacher evaluation method. One of these major needs is ensuring the individuals conducting special education teacher evaluations (i.e., principals) are unbiased and reliable (Lawson & Knollman, 2017; Semmelroth et al., 2013).

At Midwest School District (MSD), a pseudonym for a bounded system that comprised an entire school district located in the Midwestern United States, the problem of the poorly designed special education teacher evaluation process has had little focus over the years. For example, special education teachers at MSD must participate in meetings regarding students' Individual Educational Programs (IEP) and documentation of student progress. For MSD, this has recently been added in an addendum as a performance indicator for special education teacher evaluation. This addendum, however, still does not address all special education teachers' evaluation needs (MSD, 2015), which will be discussed later.

Evidence of the Problem from the Professional Literature

In local school districts and in states throughout the country, additional concerns exist that evaluation criteria for special educators are not appropriate (Williams & Dikes, 2015). Currently, special education teachers' evaluation methods are the same used for general education teachers, and therefore, they do not consider the responsibilities and other factors which these professionals must take into account (Williams & Dikes, 2015). Because of the need to provide specified support for special needs students, urban and rural school districts are implementing forms of inclusive instruction (Williams & Dikes, 2015). These changes mean that special educators are responding to the broad range of special needs students' academic, linguistic, behavioral, and social needs (Anderson, Smith, Olsen, & Algozzine, 2015). Consequently, evaluations of special educators should reflect these expanded roles (Williams & Dikes, 2015). As the NCCTQ (2010a) further pointed out, special educators address distinctly different contexts than general educators, and these distinctions are pertinent when developing an evaluation system intended for all teachers.

Although current evaluation systems concentrate on effective teaching and improved student achievement, it is important to determine how these interdependent roles may differ for special education teachers, and how evaluation systems should best reflect these differences. Teacher evaluation systems are only valid when they consider the specific responsibilities which teachers contend in the course of seeking to improve student learning and teaching (Williams & Dikes, 2015). Teacher effectiveness is at the core of education reform, according to former President Obama and his administration,

which, in turn, has created an emphasis on teacher evaluation (U.S. Department of Education (2016). An evaluation of contributions and evaluation of special educators by their districts could provide important information useful in the development of all teachers in positively supporting the achievement of students. In other words, a relationship is shown between teacher evaluation and student success (Hallinger, Heck, & Murphy, 2014; Marzano, 2012). According to more than a decade of research by Hallinger et al. (2014), teacher evaluations tend to follow a path that leads to improved student achievement. First, teacher evaluations filter out inadequate performers and provide support through feedback to create a results-oriented school culture. This helps to increase teacher effectiveness, which can have a considerable impact upon student success in the classroom and in life (Hallinger et al., 2014).

Definition of Terms

Teacher evaluation: This broad term is defined as "the formal assessment of a teacher by an administrator, conducted with the intention of drawing conclusions about his/her instructional performance for the purpose of making employment decisions" (Hallinger et al., 2014, p. 8). Teacher evaluations can be used to filter out poor quality teachers and encourage teachers to perform well (Hallinger et al., 2014).

Least restrictive environment (LRE): In order to ensure compliance with the Individuals with Disabilities Education Act of 1997 (IDEA 1997), students with disabilities must be educated alongside students without disabilities. This requirement, also known as inclusion, applies whether students with disabilities are educated in public or private schools or other facilities (Illinois State Board of Education, 2000).

Significance of the Study

In this study I described principals' perceptions of the current special education teacher evaluation process and the barriers they perceived in the existing system. I collected the information to address the disconnect or misalignment between the assessment and objectives of evaluation of special education teachers. Whereas new models for evaluating teachers are emerging, only a limited number address the unique challenges presented for assessing special education efficacy. The unique features of the special educator's effectiveness require adjustments to observation protocols and more finely tuned methods of evaluating the special education teachers' contribution to student academic growth (Darling-Hammond, 2012). As more and more states and school districts seek to address requirements of educator evaluation systems, it is important to include principal feedback on how to improve the system, as well as to understand how principals balance competing demands for fairness, accuracy, and effectiveness in measuring teacher performance.

Research Questions

In this qualitative descriptive case study of a bounded system, I examined elementary principals' perceptions of the existing evaluation procedure for special educators. Research corroborates the relationship between student achievement and teacher evaluation systems (Darling-Hammond, 2012). However, research also suggested that existing special teacher evaluation systems do not recognize and reward excellent special education teachers, remove low-performing teachers, or provide support and

development to the majority of teachers who need to improve (Jones & Brownell, 2014; Lawson & Knollman, 2017). Research questions central to this study were as follows:

1. How do elementary school principals perceive evaluation of the effectiveness of special education teachers?

2. Which key indicators do elementary principals use to observe and evaluate teaching quality?

3. Are the key indicators of teaching quality different for the evaluation of general education teachers?

4. What barriers do elementary principals encounter when evaluating special educators?

These research questions were used in an online survey method among 97 principals at MSD, the bounded system that comprised the whole district. MSD was made up of 445 principals for the 2015-2016 school year. This bounded system serves over 396,000 students at 660 schools, 484 of which were elementary schools and 176 were high schools (MSD, 2016). Approximately 80.22% of all student body are considered economically disadvantaged; 17.17% were English language learners in the 2016-2017 school year (MSD, 2016). A total of 37.7% of all students were African American; 46.5% were Hispanic, 9.9% were white, and Asian, Native American, Hawaiian or Pacific Islander, or *multiracial* made up the rest (MSD, 2016).

Review of the Literature

These sources were located mainly through Google Scholar searches using keywords including *special education assessment*, *teacher evaluation*, *educator*

evaluation systems, current evaluation methods, challenges in teacher assessment,

CCSS, Framework for Teaching, and others.

Background

On a historical basis, U.S. teacher education programs differentiated between the preparation of educators to meet the specialized content or unique student needs for instruction (Semmelroth, Johnson, & Allred, 2013). Recent federal mandates require that all teachers must acquire the skills needed to instruct students with different learning needs, including students with disabilities (King-Sears, Carran, Damman, & Arter, 2012). These requirements left many educators feeling unprepared. The challenge for principals to measure differentially prepared teacher effectiveness was exacerbated (King-Sears et., 2012**).**

Because of federal regulations, students with disabilities must have the chance to learn grade-level content (Lazarus & Reike, 2013). The Individuals with Disabilities Education Act (IDEA) of 2004 and Title I of Every Student Succeeds Act (ESSA) require students with severe disabilities participate in state assessment programs. Goldstein and Behuniak (2012) noted that effectiveness of assessment and instruction for students with special needs depends greatly on the ability of the individual teacher. Students who have significant cognitive disabilities experience processing challenges associated with short-term memory and the requirement for increased frequency of skill repetition with instructional feedback. The requirement for repetitive instruction on specific skills contrasts markedly with the breadth of content presented in the general education academic curriculum. A study of Connecticut assessment data for students with

significant cognitive disabilities showed that only half the target population could engage

with the breadth of academic content on that assessment (Goldstein & Behuniak, 2012, p.

200).

Traditionally, most special education teacher preparation programs were not

directly aligned with general education content, assessment, and proficiency standards

(Anderson et al., 2015). Likewise, general education preparation programs paid little

attention to teaching students with disabilities. Until recent decades, special education

was typically construed to mean separate education (Anderson et al., 2015). The passage

of Individuals with Disabilities Education Act (IDEA) in 1990 expanded special

education's definition to encompass all instructional settings (Murray, 2012). Similarly,

Every Student Succeeds Act (ESSA) requirement that all teachers must be effective and

certified led some schools to implement a coteaching model. A typical coteaching

arrangement involves the pairing of two teachers in one classroom (Murray, 2012). One

teacher may teach specific *content*, (such as math, for instance) while the other teacher is

a special educator. Marzano and Toth (2013) recommended that teacher evaluation

should be based on: (a) student progress that is shown by several types of performances

over time, formal, and in everyday situations; (b) assessment data collected over a period

of time with numerous performances with many different observers; (c) preparation and

planning; (d) an accurate division of skills delivered by teachers; (e) teaching and

learning strategies by which teaching may be improved through targeted support; (f)

specialized standards by which to evaluate principals and school district administrators in

their aptitudes to support and evaluate their teachers (Marzano & Toth, 2013)

Conceptual Framework

This study makes use of the Danielson framework for teaching (FFT), an important evaluation tool for teaching in the United States (Darling-Hammond, 2012). It was accepted as the single approved model in more than 20 states, including the school district of the current study. As the current study focuses on improving special education teachers' evaluation methods by determining principal perceptions toward these systems, the FFT is especially useful. According to Danielson (2010), an effective system "of teacher evaluation must answer four questions: (1) How good is good enough? (2) Good enough at what? (3) How do we know? (4) And who should decide?" (Danielson, 2010, p. 35; Evans, Wills, & Moretti, 2015). Danielson's approach was used to examine the evaluation of special educators.

The FFT has four domains of teaching responsibility that principals consider during teacher evaluations:

1. Planning and preparation comprising 25% of the Teacher Practice score

2. The classroom environment comprising 25% of the Teacher Practice score.

3. Instruction comprising 40% of the Teacher Practice score.

4. Professional responsibilities comprising 10% of the Teacher Practice score (Danielson, 2010).

The study is guided by the FFT for evaluating teacher growth and student achievement. Teacher evaluations are put in place to determine which teaching methods are successful and which are not (Goe, Hoelheide, & Miller, 2014); the FFT, therefore, theoretically helps determine special education teachers' efficacy in classrooms.

Currently, FFT is used by the MSD as the basis for considering professional development, as well as mentoring and evaluation practice. In this study I considered elementary principals' perceptions of the evaluation of special educators utilizing FFT for teaching. According to Evans et al. (2015), four special education teachers in their study argued that FFT is not always accurate for evaluating special education teachers' efficacy. Despite this argument, the FFT was used as the foundation of the school district's professional development, mentoring, and evaluation practice. These activities help teachers become reflective practitioners (Danielson, 2013; Evans et al., 2015). The research questions are based on FFT that describes the practices, skills, and characteristics that effective teachers should possess and can still hold value in this study. The FFT is used as an evaluation tool for all elementary teachers in the MSD.

FFT remains the protocol for classroom observation that a majority of states use as a rubric for teacher evaluations. The Danielson Group (2014) adapted the FFT for use in special education scenarios, in which they support special educators and their supervisors. The FFT addresses the unique characteristics of students with special needs. Illinois is one of 23 states that mandates or recommends FTT for use as the foundation for evaluation instrument (Maine Department of Education, 2012). In addition to states adopting the FFT, hundreds of districts have also adopted FFT, including Hillsborough County Public Schools, the Los Angeles Unified School District, and the Pittsburgh Public School system (Teachscape, 2011).

The New Teacher Project (TNTP; 2013a) argued that Common Core State Standards (CCSS) for teacher evaluations did not have enough updated observation tools.

This failure results in principals and other observers being unable to give specific, high-quality, and effective feedback based on classroom observations. TNTP (2013a) stated that teacher evaluations should be conducted using "grade-appropriate" CCSS; however, such standards are difficult to apply to special education teachers. While federal mandates require schools to give students with disabilities the opportunity to learn grade-level content, evaluation systems do not differentiate appropriately to address challenges special educators face. Students with disabilities are given instruction with nondisabled students; that does not mean they should be assessed on grade level content. Teachers must provide inclusive instruction for students with special needs who may lag far behind the grade level of general education students. Conducting teacher evaluations using grade-appropriate CCSS for students with disabilities is an unfair practice.

The Status of Evaluation Systems

Teacher evaluation is an important topic to discuss (Sawchuk, 2015). Various stakeholders have performed significant amounts of research that explore teachers' performance assessments. This interest is motivated by a desire to increase teacher effectiveness, by a need for policymakers to revise state laws on evaluation, and by political pressure to dismiss teachers who perform poorly. However, results from recent changes to evaluation systems remain difficult to quantify (Sawchuk, 2015).

The new system would also need to factor in student achievement. In addition, philanthropies such as the Bill and Melinda Gates Foundation spent more than $700 million on teacher quality initiatives, including creating improved teacher evaluation systems (Sawchuck, 2014). Consequently, states responded to these incentives by

rewriting laws governing teacher evaluation. Evaluation methods can include checklists from teacher observation, review by peers, portfolios, as well as FFT and value-added measures (VAM) (Benedict, Thomas, Kimerling, & Leko, 2013). Determining whether these expanded efforts produced effective evaluation systems requires further study and analysis.

Additional efforts by states and school districts include the use of statistical techniques such as VAMs that are intended to filter out sources of bias in test score growth to allow for measuring each teacher's contribution to student learning (Sawchuk, 2015). Efforts to improve standards for evaluating teachers have resulted in the growing use of VAMs. Even though a number of states use VAMs, their validity and reliability have been called into question (Murphy et al., 2013). One problem the increasing use of VAMs causes is that they shift emphasis from teachers' personal traits held by teachers, such their ability to work with others. This change may have the unintended consequence of influencing who decides to enter teaching (Harris, Ingle, & Rutledge, 2014).

The American Statistical Association (ASA; 2014) advised caution in the use of VAMs to evaluate teachers. The ASA points out VAMs do not directly measure teachers' contributions to student success but are based on standardized test scores instead. Studies using VAM indicate that teachers account for as little as 1% of test score variations, though variation as high as 14% has been shown. The ASA noted that most test score variation is due to "factors outside of the teacher's control, such as student and family background, poverty, curriculum, and unmeasured influences" (p. 7).

Similarly, the researchers in Brookings Institution (2014) called into question the validity of classroom observations for teacher evaluation. One of the key findings, was that current modes of teacher evaluation indicated vast difficulty in achieving a high rating, among teachers whose students' performance was poor. Researchers in Brookings Institution showed that teachers whose students had higher incoming achievement levels received superior average classroom observation scores than those given to teachers of students with lower achievement levels. They concluded that school districts had no processes in place to correct for this bias (Whitehurst, Chingos, & Lindquist, 2014). Assignment of students to teachers leads to systematically placing higher performing students in classrooms with higher performing teachers (Kalogrides, Loeb, & Beteille, 2013). Such assignments of students to teachers makes it difficult to determine the contributions which teachers have made to student learning.

Classroom Observations

The Bill & Melinda Gates Foundation described the teacher evaluation system as fundamentally broken. Because of this, the foundation recommended that districts and states can achieve high levels of reliability for classroom observations by implementing observer training and certification, a group of observers to audit them in an impartial manner, and more than one observation period when high stakes (performance evaluations) are involved (The Bill & Melinda Gates Foundation, 2018). Other education experts acknowledged that evaluation systems have more than one purpose, making it difficult to determine what must comprise teacher evaluations. Danielson (2012) argued that evaluations must stress improvement and accountability. Concurrently, Marzano

(2012) argued that the primary purposes of teacher evaluations are measurement and development.

Current observation systems are unfair to teachers of poorly-prepared or otherwise deficient students, and create disincentives for good teachers to steer clear of both poorly-performing students and the schools which have the greatest need. Whitehurst et al. (2015) conducted a study of four school districts across the country with enrollments ranging from 25,000 to 110,000 students. In their research, they found that only one in five educators could be evaluated based on gains in improvements to student standardized testing scores; the remainder had to be evaluated using other methods, including classroom observations (Whitehurst et al., 2015). None of the four districts were believed to have processes to address the potential for bias in scores derived from observation, which resulted from teachers being assigned stronger students than their peers (Whitehurst et al., 2015). The researchers offered the example of a teacher being assigned, either through luck or through administrative decision, to a higher-than-average group of students who were poorly prepared for academic success, possessed poor English skills, or had behavioral problems. Danielson's (2012) FFT assigns a *distinguished* rating for discussion and questioning methods only if questions posed by educators provide a consistent degree of *cognitive challenge*, along with allowing sufficient time for the student to respond, and to consider questions raised by students in the ensuing discussion (Evans et al., 2015). Under this classroom observation system, the teacher with the larger number of students who are challenging to teach will face greater obstacles to earning a distinguished rating than does the teacher in a gifted and

challenging classroom (Harris et al., 2014). Furthermore, classroom observation by principals in MSD (2011) showed that 11% of classroom observation ratings were consistently lower than ratings by trained observers, and 17% were consistently higher (Sartain, Stoelinga, & Brown, 2011). Such variability in ratings underscores the need for further research to promote teacher evaluation best practices.

Challenges to Defining Effective Teaching

Little consensus exists among teachers regarding how to define effective teaching (TNTP, 2013b). In a U.S. survey of 117 teachers representing 36 states and the 10 most populous school districts, those who responded indicated difficulty with defining ineffective teaching when a teacher works with students who are far behind their peers academically or who struggle with poverty or other problems at home. Further, 62% of respondents reported they know teachers they believe to be effective even though their students do not perform well. In response to being asked why such teachers' students may not perform well, respondents typically cited out-of-school challenges and other circumstances beyond the teacher's control. Similarly, Polikoff and Porter (2014) also noted research suggesting that state tests do not distinguish between teaching that is effective or ineffective.

Review of the Broader Problem

Issues such as the importance of educator evaluation systems, current evaluation practices, and the role of the principal in teacher evaluations were explored. Also, the issues examined were criticisms of current evaluation methods, inadequate methods for special education teacher preparation, and the role of student achievement. The topics of

value-added measures and assessment challenges for special education teachers and students were discussed. The research was limited to peer-reviewed sources that were less than five years old, which were located mainly via Google Scholar and Walden University library searches using keywords including *special education assessment*, *teacher evaluation*, *educator evaluation systems*, *current evaluation methods*, *challenges in teacher assessment,* and *principals*.

Importance of Educator Evaluation Systems

Educator evaluation has become increasingly important in recent years. Innovative teacher and principal evaluation models were developed from previous U.S. Department of Education (DoE) initiatives, including Race to The Top (RTTT), School Improvements Grants, and the Teacher Incentive Fund (Burnett, Cushing, & Bivona, 2012; Ravitch, 2016). New evaluation models can provide a basis for critical decisions concerning methods of evaluating, recruiting, retaining, developing, and compensating "human capital." When considering the importance of educator evaluation systems, it is critical to understand the models used, as this understanding offers insight into what makes a teacher successful. In the same way, the importance of educator evaluation systems at the local setting highlights what is lacking in the evaluation system in regard to special education teachers. An effective teacher evaluation program can be comprehensive and values teacher learning and growth which, in turn, can contribute to enhanced student learning (Darling-Hammond, 2015; Harris et al., 2014, Marzano, 2012). Thoughtfully implemented measures of teacher effectiveness can identify where educator practice is weakest (Danielson, 2012; Harris et al., 2014).

Over time, school systems began to implement complex observation frameworks intended to measure the value teachers added to student success. Harris et al. (2014) discussed the benefits and drawbacks of teacher observation measures typically conducted by a principal or outside evaluator using a protocol or rubric during an informal walkthrough or a formal session. Whereas some stakeholders, including teachers, principals and community members, consider classroom observations as an effective means of measuring teacher quality, researchers found that strong training and necessary "recalibration," as well as observation methods of strong validity and sufficient observation time were needed for observation to be reliable and valid (Darling-Hammond, 2015, Harris et al., 2014;). However, the current evaluation method does not work for special education teachers, as their students have various learning disabilities, disorders, and mental capacities than the students of general education teachers.

In addition, the time required for teacher observation and the collection of data, as well as to observe educators, may constitute an undue burden on school administrators (Harris et al., 2014). Further, observations may not offer teachers useful feedback if such observers are not well-versed in the content areas taught by the teachers they are observing (Darling-Hammond, 2015). Other studies indicated that administrators need extensive training to be able to make subtle distinctions between teachers who are more and less effective (Darling-Hammond, 2015).

Current Evaluation Practices

Until recent years, teacher evaluations depended almost exclusively on observations of classrooms conducted by a principal or other administrator untrained in

the process (Burnett et al., 2012). One researcher referred to these types of reviews as "traditional drive-by evaluation" (Phillips, Balan, & Manko, 2014). A typical observation consisted of a principal filling out a checklist regarding teacher behaviors and classroom attributes and failing to place the necessary emphasis upon the quality of instruction being delivered. Significant flexibility existed in how the evaluations were performed.

While teacher evaluation is required in most states and school districts, evaluation processes are centered on political goals of accountability, as opposed to practices designed to facilitate effective classroom learning and achievement (Phillips et al., 2014). Principals complained of insufficient time to conduct a thorough review, evaluation instruments with little validity, and teachers' unwillingness to change. Teachers, on the other hand, viewed the process as a meaningless exercise, concluding that for an evaluation to be effective it must provide an accurate rating, a meaningful appraisal, and an opportunity to engage in a dialogue about how to improve teacher instructional and classroom management practices (Phillips et al., 2014).

Further, teachers' unions were disinclined to support evaluation systems that could result in job loss or support the placement of a merit-based pay scale differential (Phillips et al., 2014). A valid teacher evaluation system must clearly define the standards to be evaluated and support ongoing training for the administrators who conduct the assessment, as well as for the teachers who are evaluated (Burnett et al., 2012). In other words, the evaluators must be reliable and unbiased in order to provide effective and relevant teacher evaluation scores. This topic is discussed in further depth later.

Teacher buy-in also critical to a meaningful evaluation system (White, Cowhy, Stevens, & Sporte, 2012). Teachers need to be able to take ownership of personal growth and development, and feel they should have some power over how and by whom they are evaluated (Phillips et al., 2014). Effective instruction allows students to learn in spite of their learning differences (Darling-Hammond, 2012). A special education teacher who understands how students learn and knows how to motivate them through learning difficulties is critical for teacher effectiveness in the classroom (Phillips et al., 2014).

The Role of the Principal in Teacher Evaluation

In recent years, principals have grown increasingly important in their ability to impact student achievement (Murray, 2014). As federal laws and school reform move to emphasize the general classroom inclusion of students of all abilities, principals play a significant role in promoting inclusion. Derrington (2014) studied principals' perceptions of teacher evaluation policy implementation. This concept is especially important, as this study focuses on principals' perception of special education teachers' work performance. In Derrington's (2014) study, all 14 principals and four superintendents participating in the study agreed that the new teacher observation rubric helped increase principals' knowledge and recognition of good teaching. There was 100% agreement on identifying barriers to conducting teacher evaluations. Moreover, all participants reported problems with time constraints as they struggled to implement sound time-intensive observations. Three of the four districts studied supported principals by adding personnel to evaluate teachers, thereby reducing the time principals spent evaluating. One district provided special education directors from the district office to evaluate all the district's special

education teachers (Derrington, 2014). These findings were reflected in the work by Murray (2014), whose research found that requiring evaluators to follow a set guideline or undergo additional training increased rater reliability. This concept was also present in research by Semmelroth and Johnson (2013), Jones and Brownell (2014), and Brownell and Jones (2015).

While the new evaluation system was intended to increase time spent doing classroom observations and evaluations, in reality principals spent less time in classrooms because of the demands of the new evaluation's reporting and monitoring system (Derrington, 2014; Murray, 2014). Another unintended consequence resulted from inconsistent implementation and different interpretations of state policy among districts (Derrington, 2014).

<div align="center">Criticisms of Teacher Evaluation Methods</div>

While state overhauls evaluation systems for teachers, effectiveness of those evaluation systems is inconclusive even though they are redesigned to capture and measure teacher outcomes (Smylie, 2014). According to the national survey of more than 1,000 teachers, only 25% of the teachers viewed a recent evaluation as useful and effective (Smylie, 2014). Teachers' resistance to change resulted from experience with previous innovations that were ineffectively planned or poorly supported (Harris et al., 2014). Shortcomings of previous K-12 standards over the last two decades included poor quality of the writing of content standards and developments of the assessments leading to vague understandings of content expectations across states (Harris et al., 2014; Polikoff, 2014).

Inadequate Preparation and Attrition Rates for Special Education Teachers

Implications of educator evaluation systems include their impact on special education teacher working conditions and attrition rates. Many teachers are inadequately prepared to serve students with emotional/behavioral disabilities (EBD) (Kindzierski, Marable, Odell & Raimondi, 2013). Noting the average career length of special education teachers as being only 8 years, the authors cited working conditions as causes of special education teacher attrition (Kindzierski et al., 2013). In their survey, Kindzierski et al. (2013) also found that 55% of teachers viewed the need to differentiate instruction as essential to the emotional/behavioral disabilities (EBD) classroom. Respondents also noted they themselves had observed others having difficulties in meeting instructional needs of their EBD classrooms (Kindzierski et al., 2013). Arrieta and Palladino (2015) also found that teachers offering instruction in a variety of methods typically found higher success rates among their students' understanding, comprehension, and classroom performance.

This leads to a need for future teachers to obtain multiple courses in teaching special education (Arrieta & Palladino, 2015; Kindzierski et al., 2013). More than half the teachers surveyed suggested a need for a larger number of special education preservice instruction on topics that traditional methods classes only touched upon (Arrieta & Palladino, 2015). This training should include topics such as therapeutic intervention for crises, creating and applying education plans, functional behavioral assessments, and a support group for special education teachers (Kindzierski et al., 2013). Furthermore, there is much evidence to show that there are fewer special educators than

are necessary (Berry, 2012). This shortage should further motivate efforts to assess special educators fairly, thereby increasing teacher retention (Arrieta & Palladino, 2015; Kindzierski et al., 2013). Demands of the special educator position along with the need for professional support can influence teacher attrition (Williams & Dikes, 2015).

<p style="text-align:center">The Role of Student Achievement</p>

Whereas a number of researchers advocate for the use of student achievement in educator evaluations, the role of assessment data remains controversial (Phillips et al., 2014). Standardized tests have been criticized because of the tests' narrow focus and failure to test higher-level cognitive skills. Further, the validity of using student achievement is a matter of debate due to the many elements that can affect student achievement. These elements include factors specific to the student, such as socio-economic status, mobility, availability of home support, peer culture, and prior experiences and teachers. Factors specific to the school include class size, available student support and learning resources, and particular assessment instruments used. The Institute For Modern Pedagogy And Creative Teaching (IMPACT) system, introduced in the District of Columbia Public Schools, also incorporates student achievement, along with classroom observation, teacher professionalism, and collaboration (Dee & Wyckoff, 2013). Marzano and Toth (2013) argue that evaluation must include student achievement incorporating the following:

- Student growth presented across a range of methods, as assessed both formally and in a daily capacity

- Collecting and triangulating evaluation data over time based upon multiple
  observers providing strong and frequent findings

- Adequate preparation and planning

- Ensuring that teachers hold a 'realistic' skill distribution

- Strategic improvements to provide support where it is necessary

- A model from a hierarchical perspective which evaluates principals and
  district administrators' abilities to evaluate teachers, and to provide them
  with the support that they need (Marzano & Toth, 2013, p. viii).

The above highlights mentioned to evaluate teacher success can help focus on creating an

effective teacher evaluation.

Value-Added Measures (VAM)

In some states, performance evaluations of teachers and principals must consider

using value-added measures. Laws requiring equal opportunities to access quality

education have resulted in a large number of disabled American students being taught in

general education settings (McCaffrey & Buzick, 2014). Education reform policies

require that schools and teachers need to be accountable for all students' learning or not

learning (Evans et al., 2015). Consequently, teachers are evaluated on value-added

measures intended to represent the unique contribution of the teacher to student learning.

Students with disabilities, however, present challenges for value-added calculations

(Harris et al., 2014). Disabled students tend to have low scores on regular state

assessments. Accommodations that disabled students use can affect the validity of the

score. Further, a large number of disabled students receive instruction from more than

one teacher. Students may be taught by different teachers, or they can be taught by multiple teachers in the same general education class room (Evans et al., 2015). Students with disabilities typically receive assistance from special services and aides (Derrington, 2014). As a result, determining the contribution made by each teacher is nearly impossible measurement challenge (Derrington, 2014).

Additional problems exist with value-added measures. In their small classes, special educators of students with severe disability may present with numbers too small to calculate value-added scores (McCaffrey & Buzick, 2014; Phillips et al., 2014). These and other systematic errors can be partially removed from value-added calculations by attempting to account for as many factors as possible that are related to teaching disabled students. The inclusion of disability status in the value-added model, however, may incentivize increasing the portion of poorly-achieving students who receive referrals to special education classrooms. Many students with disabilities who may require such referrals are racial minorities, or come from low-income families. While VAMs are touted as tools to mitigate the effects of standardized tests to some extent, they introduce additional problems that reduce the validity of the VAM formula.

VAM is calculated by taking numbers from standardized test scores, class size, attendance, age, disability, and English proficiency (Kourkounis, 2014). Objections were raised by teachers, school district officials and union representatives across the state, reporting that value added calculations misrepresent a teacher's effectiveness. (Kourkounis, 2014). Holdheide et al. (2012) discussed the benefits and drawbacks of evaluating special educators from a place of disabled students' achievement growth.

While VAM provides a quantitative and objective measure of teacher facilitation of their students' gains in learning over time and captures growth for all students, including those who perform at a level below proficiency levels, this model has its limitations. The challenges of VAM for students with disabilities include its inconsistent use of accommodations, mobility of some students with disabilities, and poor estimates of teacher influence regarding student performance. Consequently, states and school districts should use additional steps to increase confidence and validity of evaluation results. Another implementation issue for using VAM is that teachers may not be as likely to wish to teach disabled students. Testing data from MAP (Measures of Academic Progress) is calculated by a VAM and the test scores are linked to teacher evaluation in MSD. This is a statistical measure that takes into consideration important student variables such as student's IEP and poverty status. Teachers Union is one of many opponents of value-added measure and has advocated against the use of VAM in MSD teacher evaluation (CTU, 2016).

Assessment Challenges for Special Education

Increasing numbers of emotionally disabled (ED) students are taught part of their instruction alongside their general education counterparts (Gable, Tonelson, Sheth, Wilson, & Park, 2012). However, research indicates many special education and general and special educators do not mount the necessary preparation to efficiently apply evidence-based classroom practices required to meet ED students' academic and nonacademic needs (Evans et al., 2015). Some special education students show deficiencies in social skills and act out, are disruptive, or portray aggressive behavior. For

instance, special and general educators are often ill-prepared to teach social skills to their students (Gable et al., 2012). Students with emotional and behavior disorders (EBD) can display unusual characteristics and behaviors. These students need extensive support and resources to achieve academic success. The supports they need are student-dependent. According to (Cancio, 2013), without proper supports for these students, the outcome remains bleak. The EBD student population experiences low grades, poor social relationships, and high drop-out rate from high school (Derrington, 2014). In future, that can lead to substance abuse and unemployment. Special education in the U.S. currently has a considerable shortage of educators. The main area that needs special education teachers is for teaching EBD students (Cancio, 2013; Derrington, 2014 EBD students often have extensive needs and require pointed interventions and education by qualified professional educators. Unfortunately, teachers who work with EBD show a higher likelihood of becoming "burned out," more so than other special educators (Christensen, 2015).

Alternative Evaluation Approaches

Using standardized evaluations may compensate for aspects of an evaluation process that lack consistency or is subject to widely varying interpretations. The Teacher Advancement Program (TAP) is one such standards-based approach based on Danielson's model (Phillips et al., 2014; Evans et al., 2015). The TAP approach incorporates three aspects of teaching: instruction, the classroom or learning environment, and the planning of educational interventions (Phillips et al., 2014). The

model also uses a peer evaluation system in which teachers trained in evaluation participate in evaluating lessons and portfolios as part of the supervisory process.

Likewise, the Toledo Peer Assistance and Review Program seeks to ensure that educators hold responsibility for evaluating a dozen teachers on subject knowledge, degree of professionalism, classroom management, and teaching (Phillips et al., 2012). The principal is involved in an adjunct capacity in this process, but actual evaluations are done by the consulting teachers who have been released from the classroom specifically for the purpose of performing evaluations. Using peer evaluators has the dual advantages of reducing the burden on the principal as well as allowing for more frequent observations (Phillips et al., 2012).

States such as Rhode Island, Maryland, and New York where scores from standardized evaluations are not used or are unavailable, are considering using student learning objectives (SLOs) in teacher evaluation (Gill, English, Furgeson, & McCullough, 2014). The SLO process uses a system in which educators analyze students' performance levels and, based on that analysis, help to set appropriate year-long goals for the classroom, the school, or the individual's skills (Gill et al., 2014). Benefits of using SLOs include their applicability to all teachers in all teaching contexts, their similarity to IEPs, and their ability to capitalize on existing classroom assessments. SLO drawbacks include the possibility that disabled students may be disregarded in the SLO procedure (Gill et al., 2014). Also, SLO fails to control for factors which lie outside of the control of a given school or teacher, possibly leading to teachers resisting inclusion of disabled

students in their classrooms. Further, providing individual SLOs might reduce the necessity for educators to be responsible for all students (Gill et al., 2014).

In using the SLO process, teachers can be provided school-wide as well as other types of scores that add value to groups in nontested academic content areas and scores which are used to determine student advancement for ratings of teacher performance (Gill et al., 2014). If standardized test scores are unavailable, teachers can receive a school-wide or the group's score grounded on standardized evaluations in the academic subject (Evans et al., 2015). For example, when foreign languages are not included on a school's state standardized assessment, these teachers are evaluated and given a "score" derived from the entire schools' language arts and reading performance. Benefits of this model include the fact that the school's technological structure and approach is already in place, and these scores promote school-wide ownership of all students (Evans et al., 2015; Darling-Hammond, 2015). Challenges specific to students with disabilities include the fact that educators may be deemed responsible for the test scores of students they never taught or in any way influenced (Darling-Hammond, 2015).

Increasingly Vigorous Teacher Evaluation Methods

In the past decade, general teacher evaluation methods have undergone major changes (Coulter, 2013; Jones & Brownell, 2015). Researchers have shown that this is due to several reasons including ineffective evaluation methods. According to Coulter (2013), the previous general teacher evaluation methods used a "satisfactory or unsatisfactory evaluation" method and replaced it with a "multi-tier model with a long list of specific criteria that teachers must now meet" (p. v). Likewise, research by Jones

and Brownell (2015) found that newer evaluation methods are "more promising," as they are applicable for use in multiple "instructional settings and formats" (p. 112).

The change in teacher evaluation methods came after research indicated that the old arbitrary system was proven to be ineffective (Semmelroth & Johnson, 2014). Instead, updated teacher evaluation systems use "a variety of different evaluation methods to measure teacher effectiveness"; still, these evaluation methods are subject-specific for math and reading (Semmelroth & Johnson, 2014, p. 131). New teacher evaluation methods require teachers to meet specific criteria include the recognition of needs held by students, reliable focus upon educational content, use of data derived from student observation to drive instruction, and strong expectations, among others (Coulter, 2013). As stated above, these new criteria for teachers to meet replace the old "satisfactory" or "unsatisfactory" teacher evaluation systems, allowing more room for feedback and improvement on specific areas of teaching instruction (Coulter, 2013; Jones & Brownell, 2015; Semmelroth & Johnson, 2014).

Despite these improvements to the general teacher evaluation systems, however, progress still must be made in developing strong evaluative methods for special educators, according to Semmelroth and Johnson (2014). Specifically, Semmelroth and Johnson (2014) stated that the current, updated teacher evaluation methods are geared toward specific areas of content and ability. In other words, no adequate special education teacher evaluation system exists. Semmelroth and Johnson (2014), therefore, argued: "Special education teachers work in highly specific but diverse instruction environments," which requires a "wide range of roles and responsibilities" (p. 132). The

authors stated that research has proven that current observation "has not sufficiently considered special education," and imply that there is a need to create effective special teacher evaluation methods (Semmelroth & Johnson, 2014, p. 132). These findings are echoed in studies by Coulter (2013), Jones and Brownell (2014), and Lawson and Knollman (2017).

Increased Training for Special Educator Evaluation

The research indicates that current teacher evaluation systems fail to adequately consider the special education classroom environment (Lawson & Knollman, 2017; Semmelroth & Johnson, 2014). While evaluations for general education teachers have evolved over the past decade (Coulter, 2013; Jones & Brownell, 2015), these same evaluation systems do not apply well to special education teachers' many responsibilities (Lawson & Knollman, 2017; Semmelroth & Johnson, 2014). This creates a need for research to be focused more on creating a relevant evaluation system for special educators instead of continuing to focus on improving the existing system for general education teachers. In fact, although policy initiatives have promoted "comprehensive and rigorous evaluations of teachers," these methods fail to adequately measure "effective teaching," especially for special education teachers (Lawson & Knollman, 2017, p. 6).

Corresponding research by Jones and Brownell (2014) found that the continued focused efforts on improving general education teachers' evaluation methods has left a lack of interest or willpower in improving special educator evaluation systems. According to Jones and Brownell (2014), the "research methods are yielding important

information about the effectiveness of general education teachers, particularly those

providing language arts and mathematics instruction"; however, because these studies

have not included special education instructors, it has left "states and districts grappling

with how to adjust evaluation systems to deal with the unique needs of these teachers" (p.

112). Furthermore, the researchers stated:

> The lack of attention to special education teachers in this research is likely the
>
> result of unique challenges associated with measuring teaching effectiveness in
>
> special education, including special educators sharing instructional
>
> responsibilities with other teachers, variation in student ability levels, and special
>
> education teaching occurring across multiple settings, (e.g., self-contained
>
> classrooms, resource rooms, and in coteaching arrangements in general education
>
> classrooms). (Jones & Brownell, 2014, p. 112)

As stated, there are various unmet needs of the special education teacher

evaluation method. One of these major needs is ensuring the individuals conducting

special education teacher evaluations (i.e., principals) are unbiased and reliable,

according to Lawson and Knollman (2017) and Semmelroth and Johnson (2014).

Semmelroth and Johnson (2014) found that employing many reviewers in various

settings (including 'explicit' instruction and group lecture) is the only way to ensure

reliability and validity within special education teacher evaluation systems. These

findings align with those which found that "multiple observations and multiple raters are

critical for ensuring acceptable levels of measurement score reliability" (Semmelroth &

Johnson, 2014, p. 131). Additional studies by Lawson and Knollman (2017) determined

that although administrators may feel that they have the "ability to provide fair and meaningful evaluations of special education teachers," increased training is necessary (p. 6).

Furthermore, research by Jones and Brownell (2014) calls into question whether the current method of in-classroom observation is useful, reliable, or accurate for evaluating special education teachers. The researchers state that even today, "neither researchers nor practitioners have arrived at a consensus on the best methods for evaluating special educators" (Jones & Brownell, 2014, p. 112). Current systems like scores based on value-added elements are not useful for special educators, and while observation is shown to be superior, it must employ many raters to participate in the evaluation process, as indicated by Semmelroth and Johnson (2014).

Therefore, the most promising methods of special education teacher evaluation are observation, but only if multiple administrators take part in the observation process (Lawson & Knollman, 2017; Semmelroth & Johnson, 2014). This is because observation systems can be employed across many instructional environments (Jones & Brownell, 2014) and because more evaluation raters decrease the risk of bias and lack of compassion (Lawson & Knollman, 2017; Semmelroth & Johnson, 2014).

Implications

Implementing effective measures to evaluate special education teachers requires extensive analysis, planning, and preparation. Understanding the challenges elementary principals perceive in evaluating special educators will help inform educator evaluation practices. Using a specific observation protocol, the current study helped identify the

needs of elementary principals and administrators by which they can more effectively evaluate special education teachers. The study offers recommendations which will inform policy by which problems connected with evaluation of special educators and provides a guide for the MSD responsible for improving teacher evaluation.

Summary

This section introduced the problems of evaluating special education teachers, as required by federal legislation promoting inclusion and accountability for student growth. In addition to examining the impact of this problem on local school districts and states across the country, this section outlined a concept-driven framework for this study, as well as reviewed evaluation practices and their shortcomings. Section 2 will discuss the methodology of this qualitative descriptive case study of a bounded system, and will provide justification for the study design, before discussing goals and limitations for this exploration of perceptions held by principals of a local elementary school.

Section 2: The Methodology

Introduction

In Section 2, I describe the qualitative methods research design and approach and review participant recruitment and protection. In this section I also discuss processes informing the collection of data, and analytical methods, along with limitations of the study. This section will also include the results of my data analysis.

Qualitative Research Design and Approach

The selection of a research approach is driven by the nature of the problem under investigation (Creswell, 2014). Further, specific types of social research problems lend themselves to specific approaches. Creswell (2014) argued that if the research problem requires the identification of underlying factors that influence an outcome, a qualitative approach is best. In the context of the present study, problems associated with special education teacher evaluation need identification in order to correctly assess their impact on students, teachers, school districts, and other stakeholders (Maxwell, 2013). Creswell likewise noted that when a research problem requires understanding the utility of an intervention, then using a qualitative methods approach is most appropriate. With respect to the assessment of special educators, various alternative methods have been attempted or proposed, without clearly establishing best practices (Maxwell, 2013). Employing a qualitative method approach furthers analysis of the efficacy of current evaluation methods.

Research studies that employ qualitative methodology are especially well suited to generating various factors that explain insights (Maxwell, 2013). It is believed that

perceptions held by principals regarding performance evaluation for special education teachers can lead to revelations that advance understanding of the effectiveness of evaluation practices. Given that qualitative research seeks to explain extensive details, surveys of elementary principals' perceptions of special educator evaluation are suited to providing relevant information. Qualitative research therefore provides the best approach to meeting the goals of this research project.

A descriptive and qualitative case study was employed, which used a process of preliminary data analysis. The data were collected from closed and open-ended survey questions, where 'closed' questions were answered with simple responses, and open-ended questions solicited more expansive responses from subjects considered. For the open-ended questions, written responses were checked and tracked. Associations between emergent issues were accumulated into potential themes to analyze the coded data. The survey focused on elementary principals' perceptions of special education teacher performance reviews. In addition to closed survey items, the instrument also includes open-ended items (Coryn, Noakes, Westine, & Schroter, 2011). Maxwell (2013, p. 31) cited the example of a study using open-ended items on a questionnaire system as having much greater insights with the school administration.

Formative evaluations are intended to provide a foundation for improvement by identifying strengths and weaknesses of a program or process (Coryn et al., 2011). Summative evaluation on the other hand compares the current assessment process against some standard or benchmark not found in this study sample. In addition, this research does not seek to clarify which outcomes have been met since these are established by

state and federal mandate; outcomes-based evaluation does not fit this study's purpose. Similarly, this study does not attempt to measure specific goals in special educator assessment, and so does not consider how to measure progress toward teacher evaluation goals. Instead, the overall goals were to understand the perceptions of elementary principals regarding their evaluation of special education teachers, the indicators they look for, and the barriers they encounter during the evaluation process.

Participants

The study's setting was the Midwest School District (MSD), a bounded system that comprised the whole district. MSD was used as a pseudonym to protect participants' identities and to not disclose the school district being studied, though it is reasonable to disclose that it is located in the central United States.

MSD was made up of 445 principals for the 2015-2016 school year. This bounded system serves over 396,000 students at 660 schools, 484 of which were elementary schools and 176 were high schools (MSD, 2016). Approximately 80.22% of all student body is considered economically disadvantaged; 17.17% were English language learners in the 2016-2017 school year (MSD, 2016). A total of 37.7% of all students were African American; 46.5% were Hispanic, 9.9% were white, and Asian, Native American, Hawaiian or Pacific Islander, and those who were multiracial comprised the rest (MSD, 2016).

All MSD elementary principals ($N = 445$) were contacted using directory information. Ninety-seven elementary principals responded to the survey. Sample inclusion criteria were elementary school principals currently employed by the MSD. The

district consisted of 445 elementary schools and elementary principals. The number of elementary principals responding to the survey was 97 or 22% of the elementary principals reported as of September 2015 (MSD, 2015). I used Survey Monkey to send an email to all 445 principals with an invitation to participate in this survey during the summer before the 2015-2016 school year. The study setting was the MSD and an online web-based survey platform was employed to disseminate the survey and to collect survey responses.

Most of the principals were female (59.0%), within the age range: 46 to 50 years (23.7%), highest level of education: master's degree (79%), have a teaching degree (59.8%), have a current school enrollment of 1-499 students (25.8%), and represent regional location: North/Northwest Side (32.8%) (Table 4).

By asking survey questions, it was believed that the research could yield information not anticipated. Principals identified barriers to effective evaluations and explored their perceptions of alternative assessment methods or other related topics.

Instrumentation and Materials

This study's purpose was to determine how elementary principal perceptions of special educator evaluation processes. Research questions focused on how elementary principals perceived performance evaluations of special educators, as opposed to general educators. In addition, key indicators that reflect special educators' performance were identified. Together, the research questions focused on identifying key indicators for good performance among special and general education teachers and organizational climate factors that may provide barriers to elementary principals attempting to evaluate

teacher performance. An example of questions which were employed in the survey used for this study is: It is possible to fully identify the influences of a special education students' other teachers using teacher ratings? Likert-style responses were solicited from the subjects, ranging from "Strongly Disagree" to "Strongly Agree." The survey is presented in full in Appendix B.

I developed the instrumentation based upon data from related studies, including the studies by Sledge and Pazey (2013) and Phillips et al. (2014). These studies provided a strong basis of understanding with respect to the development of an instrumental framework for the evaluation of educator performance, and different paths through which such frameworks can be established. After the instrument was written it was provided to nonparticipant principals to review for clarity and unforeseen errors.

Qualitative research focuses on codifying narratively-derived data from survey questions without particular answers, as well as other comments principals wrote to explain their responses to Likert- item statements. This research work analyzed responses to Likert-type items which measured subjects' degree of agreement with statements related to teacher performance evaluations with respect to general and special educators. Through qualitatively-focused data collection, I gathered and analyzed this information to increase understanding of participants' perceptions of elements considered in this work.

The first open-ended question asked administrators to identify goals they wanted to address in special educator evaluation process. The second and third open-ended questions asked principals about respective strengths and weaknesses of their current educator evaluation process. The fourth question asked principals what changes, if any,

they would like to make to their special educator evaluation process. The final open-ended question asked administrators to provide any additional comments on special educator evaluation practices and related issues. In addition, respondents were able to make additional comments in the comment field providing depth and detail for any of the 23 Likert scale structured questions that ask them to rate their agreement with survey items. The closed ended responses were: *Strongly Agree, Agree, Disagree, or Strongly Disagree*. The following section will consider strategies employed in data collection.

Data Collection

An online web-based survey was employed both to apply the survey and to collect subject responses. All elementary principals within the MSD were invited to participate, as were all elementary principals within the district, in order to capture data addressing the complexity of the evaluation process. The survey was administered via Survey Monkey and included elementary principals listed in a directory of MSD elementary principals for the 2015-2016 school year.

The data were collected and analyzed. The components provided in-depth insights into the participants' perceptions and gave context to the findings. In addition to collecting data on special educator evaluation practices, the survey collected, anonymously, elementary principals' demographic data, including age, gender, education status, and years of work experience as an educator and as a special educator. The demographic data of elementary principals were used to describe the participating sample.

Personal and private data were protected by Survey Monkey, which disabled IP address tracking, and used secure transmission to protect the exchange of private data. A link in the email invitation took participants directly to the informed consent form stating the purpose of the study, as well as benefits and risks they might incur by their participation, and other information necessary for their informed consent. The participants' consent to have their data collected was labeled as question 1 on the survey, to ensure respondents' attention to the conditions of the study. Reminders were sent to the participants weekly to increase the response rate. After the participant consent, he or she clicked next to begin answering the online survey.

Data Analysis

I focused on locating factors that contributed to how elementary principals evaluate special education teachers' performance. I examined all responses (open and close ended) for consistency and breadth of perceptions. I did not find any discrepant cases that would need to be eliminated from the analysis.

I summarized the data derived from the closed-ended Likert items. For responses to open-ended prompts, I used Microsoft Word to track descriptive themes, by which these responses could be interpreted and summarized effectively. Another special education teacher independently coded and interpreted responses as well, thereby increasing accuracy and validity. I compared the various coding-derived categories and performed two review cycles to refine codes, as well as categories and 'subcodes' which were derived. This coding method was hierarchical in nature, and resulted in a host of themes, from which patterns were derived.

Using an inductive approach to data analysis, I looked for patterns to identify themes emerging from responses. Analysis of open-ended responses and comments revealed several major themes consistently associated with research questions the study tried to answer. The eight thematic categories were identified and are as follows:

- The need for differentiated evaluations to reflect the nature of the special educator's role

- The need to account for skills related to inclusion and accommodation for diverse learners

- The need to support student growth in academic and social dimensions

- The need to employ measurements that are valid and reliable

- The need to validate good instructional practice

- The impact of resource constraints

- The impact of the administrators' qualifications on the special educator evaluation process

- The identification of barriers to effective evaluation

- Ways for the evaluation to address accountability for student learning outcomes

Data obtained from open-ended questions were analyzed by coding and through the use of labels and tags from which themes were derived, then used to connect meanings, themes, and categories (Van Lint, 2012). I used all responses that were sent in exactly as they appeared. All the responses were written by the participants.

The responses to the open-ended questions were used exactly as they appeared in the participants' own written words.

Limitations

The study had several assumptions. First, because it would take considerable time and effort to validate the answers provided by each subject, it was assumed that participants answered honestly, especially with respect to their professional qualifications. It was also assumed that the inclusion criterion to select school district elementary principals was appropriate for studying the research questions. Although this survey research was best suited to collecting data on elementary principals' perceptions of teacher evaluations, some limitations of qualitative research remained. One strong limitation regarded the number of choices under the Likert-type instrument, which might have been insufficient. This limitation was addressed by the inclusion of open-ended survey items. The inclusion of open-ended questions let subjects express their opinions and introduce additional response themes which could be subject to further analysis.

Bias by self-selection was seen by allowing principals to personally seek to take part in this data-collection process. Inviting the entire population of MSD elementary principals to participate in the study was an attempt to curb the bias.

This study examined elementary principals' perceptions of the methods by which elementary teachers are evaluated including both general and special education. Current evaluation methods were written to evaluate elementary education teachers, including regular and special education teachers. When special educators are assessed under similar

premises as general educators, special education teachers may be at a disadvantage because the students' disabilities could limit the teacher performance.

Delimitations are defined as the choices that the researcher assembles for the study that are under the control of the researcher (Creswell, 2014). Delimitations include the researcher's decision to use a Likert scale survey to conduct the research. More material and data may have been harvested if I had chosen to conduct interviews (either group or one-on-one interviews) or on-the-job observations. These delimitations are presented due to this study's focus upon interacting with participants in an online setting, through use of surveys. I selected elementary principals specifically for this study because the roles of elementary principals are different than the high school principals.

In particular, elementary school students face considerably different educational standards, as well as social and developmental realities, than their older peers in the high school environment. As a result, it is critical for special educators to be able to reach these students in an effective manner before they undergo the significant life and developmental changes which occur in higher grades, and become subject to more stringent educational standards and major life obstacles. Because elementary special educators must be strongly-prepared to provide special education students with adequate education and guidance, their oversight must be as comprehensive as possible. In essence, without adequate special education (as informed by principal-evaluators with strong expertise as to the specific needs of these students, and the responsibilities held by their educators), students with special needs may face severe setbacks in later years, both

in the educational environment, and in their lives in general. This district was chosen in particular primarily for reasons of convenience, as it is local to this researcher.

Protection of Participants' Rights

Permission to conduct the study was given by Walden IRB and the school district where MSD is situated. Measures that were taken to protect participants' rights included educating them with respect to the study's purpose, disclosing potential benefits to the individual and others, and disclosing the conditions of participation, including the right to refuse or withdraw without penalty. Participants were asked to acknowledge informed consent before proceeding to the survey using the link provided by the email invitation. Participants were also informed that Survey Monkey safeguards their personally identifiable private information, including email addresses. Survey Monkey holds data securely on servers located in the United States. Participants were also informed that the risk of research-related injury, including physical, psychological, social, or financial risk was negligible.

Data Analysis Results

The purpose of this study was to investigate how elementary school principals perceived the process of evaluating special education teachers of the MSD. This qualitative descriptive case study consisted of a 39-item scale using closed- and open-ended survey questions to analyze the relationship between responses within a bounded system. Twenty-three structured survey questions allowed respondents to add any details to their responses in the comment box. Some respondents provided narrative comments for each item, which added insight and depth to the closed-end response items.

All the surveys that were returned were complete and were used for analysis. Survey responses were deemed to be valid and no discrepant cases were identified.

Demographic Information

This section presents demographic data for survey respondents and profiles of their schools. The population consisted of 445 MSD principals ($N = 445$), to whom email invitations were distributed using the Survey Monkey platform. The survey remained open from June through September 2016. Second and third email reminders were sent during the summer break to increase the response rate. Of the 97 respondents who opened the survey, a smaller number completed the remaining 39 questions. The sample size for each closed-ended question ranged between 59 and 63 respondents. 59% of subjects were women, and 41% were men ($N = 61$). When asked the highest level of formal education completed, responses were ($N = 63$): 79.4% had Master's degree (including holders of multiple master's degrees), 17.5% had doctoral degree or equivalent and 3.2% had bachelor's degree. Almost 25% of respondents were between ages 46 to 50, with the remainder spread over other age groupings ($N = 60$).

Demographics Using Descriptive Statistics

Descriptive statistics revealed a sample size of 97 principals. Most of these were female (59.0%), within the age range: 46 to 50 years (23.7%), highest level of education: Master's degree (79%), have a teaching degree (59.8%), have a current school enrollment of 1-499 students (25.8%), and represent regional location: North/Northwest Side (32.8%) (Table 1).

Almost 55% of participants had not worked as a special educator, while 45% had experience as a special educator or in a related field for periods ranging from one to more than 20 years. During the 2015-2016 school year, almost 80% of the schools represented ($N = 62$) had enrollments of 999 students or less. More than two-thirds of the schools represented had enrollments of special needs students greater than 11%. A total of 59% of teachers are special educators in each of the schools represented ($N = 59$). See Table 1.

Table 1

Descriptive Statistics

Category			*%*
Gender			
	Female	36	59.0
	Male	25	41.0
Age			
	26-30 years	3	5.1
	31-35 years	8	13.6
	36-40 years	7	11.9
	41-45 years	8	13.6
	46-50 years	14	23.7
	51-55 years	9	15.3
	56-60 years	0	0
	61-65 years	0	0
	66 or older	4	6.8
Highest Level of Education			
	Bachelor's degree	2	3.2
	Master's degree	50	79.4
	Doctorate	11	17.5
Have a teaching degree?			
	Yes	58	59.8
	No	1	1.0
Current enrollment (students)			
	1-499	25	25.8
	500-999	24	24.7
	1000-1500	10	10.3
	1501-1999	1	1
	2000 or more	2	2.1
Regional Location			
	Central	12	19.7
	Far South Side	4	6.6
	North/Northwest Side	20	32.8
	SouthSide	7	11.5
	Southwest Side	12	19.7
	West Side	6	9.8

Report of Survey Qualitative Responses

This study sought to understand elementary principals' perceptions of the special educator evaluation process; their detailed comments regarding improvements to teacher evaluation are reported in the following tables and paragraphs which provided a wealth of insights. The responses to the open-ended questions were used exactly as they appeared in the participants' own written words.

Goals and Improvements for Special Education Teacher Evaluation

In the survey, when asked what goals their SPED (special education) evaluation process should address, administrators provided responses suggesting their current process either does not address special educator goals or does so only inadequately. All of the tables reported presented below includes descriptive responses in its entirety. Thirty percent of administrators used the phrase "one size fits all" to describe current ineffective evaluations. Seventy percent of elementary principals do not feel current evaluation methods sufficiently differentiate the special educator's role, as expressed in this quote: "I would like to see goals directly targeting special education teachers." Participants cited the need for a separate evaluation track or modifications to the current rubric based on Danielson's Framework. As expected, administrators believe evaluations should reflect teachers' achievements in working with students with a range of disabilities. When asked "What changes if any would you make to your special educator evaluation process?" (Q39), participants' responses included (1) accounting for the range of disabilities and creating a fair rubric that better determines what the special education teacher is teaching for better accuracy, (2) finding methods to "specifically and

effectively evaluate" these teachers based on the range of disabilities of their students, and (3) finding a system that does not require the student's disability to be specified in order to teach to them. One participant noted that "inclusion teachers don't have the same planning roles" as general educators, which affects how the evaluator uses the grading rubric. Another participant noted:

> Taking into consideration the vast range of DL (diverse learners) needs and ability levels, setting an expectation for all DLs to perform at grade level contradicts the nature of why special education exists in the first place. For some DLs, this can be an attainable goal. For others, it may not. Therefore, there is a great need for expectations to truly be differentiated with no underlying or overarching expectation that is impractical and contradictory.

The Weaknesses of Current Special Educator Evaluation Process

Not all principals felt their current evaluation process has shortcomings. One principal specifically cited a need for "rubric descriptors geared to SPED teachers." Another noted, "I think the Danielson rubric largely works for SPED teachers." One principal's comments reflected the nuances of special education. When asked to rate agreement with the statement that students with special needs students should achieve at a rate commensurate with their general education peers of the same grade level, one principal responded:

> This is difficult to answer. We need to set expectations high, but keep in mind that disabilities are very unique and personal. It is not a one size fits all model. I am a strong believer in a growth mindset and that all children should make growth;

however not all children are able to follow the same growth trajectory, which is

why they most likely have an IEP to begin with.

Principals' comments cited the following weaknesses: (1) "explicit information regarding

special education," (2) the evaluation rubrics are "not aligned to the instructional

programs that … special education teachers are following," which leads to a "lack of

understanding of student growth," and (3) the general rubric does not translate well to

evaluate special educators, as "it does not necessarily correspond with the specific

disabilities they are trying to teach." An example of this, as stated by the study

participant, is that "If a student is nonverbal, it is difficult to evaluate for student

discourse etc. (in the distinguished category)."

Desired Changes to the Special Educator Evaluation Process

Other administrators listed specific improvements to their evaluation process, in

particular changing the evaluation rubric. Suggestions range from adding an additional

evaluation rubric for coteachers to use as an addendum to the current Framework, adding

a specific rubric developed specifically for special educators. While principals offered a

range of proposed solutions, there was general agreement that improvements to the

current process are necessary. When participants were asked what changes, if any, they

would make to their school's special education teacher evaluation process, some

comments included: "Take student performance out of the criteria for Special Ed.

teachers." Many answers centered on a theme of changing the rubric for these teachers'

evaluations, as "They require a different rubric, not just an addendum." One participant

noted: "This is a bigger conversation that cannot be addressed here. But in short, a new

rubric needs to be developed that accounts for different ranges of DL ability levels and educational setting."

Need for Differentiated Evaluation for Special Educators

Whereas 70% of the administrators' comments showed a belief in differentiated evaluations for SPED (special education) teachers, this was not a unanimous point of agreement. Several principals (30%) expressed the belief that evaluations should be the same for general and special education teachers. One principal noted that in other than severe clinical cases, "The special education teachers should follow the same guidelines such as Common Core State Standards and grade level objectives." Another commented that "SPED teachers should know and practice the same strategies as regular ed. teachers as differentiated instruction is a must for both regular and special ed. teaching." When asked what changes they would make to differentiated evaluation, participants noted: "With practice and differentiated instruction, I expect students to achieve at grade level," "A modified and aligned CCSS curriculum should facilitate the expectations of equality," and "Students can be expected to achieve the same goal, but not be held to the same standards as the general education students."

Desired Goals for the Special Educator Evaluation Process to Address

Administrators have a range of responsibilities, including ensuring that they accommodate and be willing to modify their methods to meet the needs of diverse learners. Disabled students may require a change in curricula or expectations, or a change that assists the student in overcoming or working around the disability. In some circumstances, special education students may be expected to master the same material as

fully as general education students. Regardless of which assessment tool or process an administrator uses, the special educator is affected. Principals' responses discussed concerns with how SPED evaluation needs to address the need for inclusion and accommodation. When asked about goals for the SPED evaluation process to address, administrators acknowledged that adapting tailored instruction to address disabled students' particular needs was critical. Principals believe they must evaluate educators' ability to provide different types of support for individualized instruction that meet disabled students' needs. A majority of elementary principals' responses showed a belief that a special education assessment process needs to reflect differences from general education evaluations. When asked what goals they would like their special educator evaluation process to address, participants discussed inclusion versus pull out models, setting specific goals for individual students based on ability, and the possibility of the evaluation to reflect the "diverse nature of teaching diverse learner students." As one participant noted: What the teacher faces in instructing these students is not one size fit all and their progress cannot rest on one test." Other ideas included teaching strategies that meet students at their instructional ability levels.

Weaknesses of the Special Educator Evaluation Process

Another theme which emerged from elementary principals' comments concerned the need for an evaluation approach that effectively addresses student growth. When asked to discuss weaknesses of their special educator evaluation process, administrators cited a number of factors. They noted limited opportunities to look at student work products, which therefore yields a less than accurate measure of student progress and

teacher effectiveness. The goal of evaluation is measuring student growth in multiple dimensions, including social, emotional, academic, and functional goals, and the assessment process must address these goals. Principals' comments about student growth and teacher evaluation show they believe better ways of measuring teacher performance exist than what many of them are currently using. When asked "What are the weaknesses of your special educator evaluation process?" the principals cited some of the following as weaknesses affecting their ability to use the evaluation process as a way to promote student growth: Answers ranged from "It lacks the ability to accurately evaluate the teacher in meeting the needs of their students" to "Does not include a student growth metric for diverse learners into the school's SQRP" and "Parents wondering who will have the best interest of their children at hand in the classrooms." Other answers focused on the one-size-fits-all rubric that is used to evaluate different teachers, which makes it lack the ability to "accurately evaluate the teacher in meeting the needs of their students," and the growth metrics of diverse learners served by the teachers.

Comments Regarding Indicators of Teaching Quality

A research question the study sought to answer was determining which key indicators principals use to observe teaching quality. When asked what goals they wanted their special education process to address, principals cited a need for more effective measurement and increased accuracy in measurement. Perhaps reflecting the management philosophy that one cannot manage what one cannot measure, principals wanted better measurements. One administrator noted that the current process ensures compliance with statutory guidelines but was less certain of "evaluation fidelity." They

noted weaknesses in current measurement systems, including the limitations imposed by infrequent classroom observations. One principal cited misalignment between the Danielson Framework and IEP goals. When asked if assessment tools for special educator evaluation are inadequate, one principal responded "What assessment tools?" Taken together, their comments indicate that principals would prefer more accurate means of measuring indicators of teaching quality than are currently available. In terms of indicators of teaching quality, participants noted the following: they want more than test scores as the only indicator of measuring the performance of special education teachers; they want a tool that will "Account for the varying degrees of disabilities and create a fair rubric that will capture what the special educator is actually teaching for better accuracy"; and they state that the assessments data cannot be used to accurately evaluate teachers' performance. Furthermore, participants sought "a process that accurately measures the impact the teacher has on the whole child," one that "[includes] student data with general education data." Participants noted that the evaluation tool should be something that can be globally applied, not just a "one-shot observation." Others said that "The 40-minute observation and rubric does not capture their ability to teach special needs students."

Goals for the Special Educator Evaluation Process to Address

Elementary principals were asked to comment on goals for their SPED (special education) evaluation process to address or to identify changes they would like to make. Their responses discussed improvements in validating good instructional and professional practices. One administrator wanted to see, "Proper preparation on the part of

administrators to gather background information that will frame observations for evaluation of special educators are also included and given a hand in their evaluation during pre and post conference discussions."

Several elementary principals would like more coteaching as a way of accomplishing multiple goals of inclusion and educating special needs students in a least-restrictive environment (LRE). Coteaching allows students with disabilities to access the general education environment, while still allowing them the benefits of specialized instruction. Principals' suggestions did not specify a preference for a particular coteaching model, possibly indicating limited exposure to other models than the one currently employed in their school. In addition, even though coteaching produces benefits it also comes with assessment challenges. Teacher effects are still more difficult to identify using the coteaching service delivery model. When asked what goals they would like their special educator evaluation process to address, elementary principals' comments included the following: "More effective coteaching, how to support more students in being cotaught" and "Coteaching and allowing evaluations to take place for a gen ed and coteacher during the same observation, holding pre and post conferences with both together."

Teaching Special Needs Students Requires More Than Just Good Instructional Practice

Elementary principals were asked to rate the degree to which they agreed with the statement "teaching special needs students requires more than just good instructional practice." Their comments cited the need for additional practices and strategies, including

evaluations that take into account an understanding of how specific conditions or disabilities affects student brain functions and ability to process information.

Barriers to Evaluate Special Educators

Another recurring theme concerned barriers to effective evaluations. Research question 4 explored barriers principals encounter in evaluating SPED (special education) teachers. One of the most significant barriers was resource constraints. Not surprisingly, principal's perceptions of SPED evaluation included criticism of the impact of limited resources on their ability to effectively evaluate SPED teachers. Principals criticized current evaluation methods for being "tedious" and "time-consuming," cutting into the amount of time left to meet their other responsibilities.

Principals frequently cited inadequate funding as an impediment to the evaluation process. Others felt that funding was incorrectly aligned with resources actually required; there was general agreement among administrators that special education programs are underfunded. Elementary principals want to change or perceive weaknesses of the current evaluation process through many ways, including more time ("Not enough time to observe for a higher frequency of visits"; "Less compliance and administrative responsibilities for principals, so "instructional leadership" can have more time"; "Longer chunks of time, dedicated sub so that we can pull teachers out for longer or money so that we can pay them to stay after school to review case studies") and financial support (which should be "based on number of minutes, not the number of special education students"; other comments included "budget seems to trump student needs" and "This

area has been deprived of funding for many years. We need our Special Needs students to have placement where they can have strong goals and achieve their potential").

Desired Changes to Make to the Special Educator Evaluation Process

Even though elementary principals believe their current evaluation process is largely effective, they see there is room for improvement. The following comments indicate specific improvements respondents believe would improve the evaluation process. They indicate support for the questions this study explored: differentiated special educator evaluation, use of appropriate evaluation measures, differentiated key indicators of teaching quality for special educators, and the existence of barriers to the evaluation process. When asked "What changes, if any, would you make to your special educator evaluation process?" participants' responses included to take the student performance out of the criteria for special education teachers and to have an experienced special education administrator evaluate special education teachers. Notably, other participants stated: "a new rubric needs to be developed that accounts for different ranges of DL ability levels and educational setting" and "Account for the varying degrees of disabilities and create a fair rubric that will capture what the special educator is actually teaching for better accuracy." Participants wanted a process that "accurately measures the impact the teacher has on the whole child"; they specified that "the 40-minute observation and rubric does not capture their ability to teach special needs students."

Administrator Qualifications to Evaluate Special Educators

The study also sought to understand how elementary principals perceived adequacy of their own qualifications to effectively assess special educator performance.

Principals were asked to demonstrate the degree to which they agreed with the statement that they did not "have prerequisites to conduct special education teacher evaluations." The majority of principals expressed a strong belief that they had the necessary background. A later question (question 31) found that out of 62 who responded to this question, majority of participants ($N = 34$) however have had zero years' experience working as a special educator. When asked about their own competence and skills related to evaluating special education teachers, while most participants noted that it would only be fair that an individual who is experienced in dealing with special education students should be allowed to evaluate special education teachers' performance

Similarly, 54% of participants agreed that the precise evaluation of a teacher's efficacy might be reduced when the evaluator lacks a basis of awareness of specific practices from which student outcome improvements are derived. This topic was also touched on in question 18, "Teacher evaluation is objective when the evaluator has experience in teaching and assessing students with special needs," of which 28 participants agreed. (However, interestingly, a large number, 21 participants, also disagreed.) These findings were underscored in question 4, "Teaching special needs students requires more than just good instructional practice," with which the majority of participants ($N = 29$) agreed. Furthermore, the majority of principals ($N = 32$) stated that they "strongly agreed" with the notion that teaching special needs students requires instructional strategies teaching social-emotional skills. These skills are often different from those used to teach general education students.

More than one administrator noted that having experience as a special educator, school counselor, or case manager strengthened their evaluation process. Another saw particular value in having an experienced special education administrator evaluate special education teachers. Another administrator noted that even if he had not had special education experience, the Danielson Framework work have enabled him to conduct evaluations. Many stated that they do have the prerequisites to conduct special education teacher evaluations. In contrast, one principal noted feeling unprepared to evaluate the related service providers (speech pathologists, psychologists, social workers, etc.).

Strengths of the Special Educator Evaluation Process-Accountability

Another important theme related to evaluation was the concept of accountability. Respondents felt it was important to demonstrate standards that all teachers were held accountable including special educators. When discussing the strengths of their special education teacher evaluation process, participants noted that it was equal to the general education teachers' evaluation process and that all teachers are held accountable and their expectations are set high. One participant noted: "There is an evaluation process and some level of accountability. That is important."

Additional Comments

Exploring elementary principals' perceptions of approaches they use to evaluate special education teachers revealed aspects of teacher effectiveness that cannot be explicitly measured. When asked about the strengths of their special educator evaluation, principals made comments such as the following:

- "It is just a snap shot. Ability to pull together an awesome Special Olympic program for students and their families is not weighted but is a valued skill in a sped teacher";

- "They love the kids and their jobs";

- "How do you give credit to the teacher who has the intangibles of excellent rapport and the persistence to have students be successful vs. the teacher who has a similar skill set but does not exhibit the same passion or drive?"

Another administrator noted one of the disheartening aspects of evaluation systems that continue in a state of flux: "Teachers are not sure from year to year what the formula will be and if they have their jobs."

In reviewing these themes and others, it is clear that principals perceive varying levels of success resulting from the current evaluation process. At the same time several principals acknowledged the challenges of special educator evaluations, they also noted their belief they were effectively meeting those challenges, as shown by their current process. When asked to cite the strengths of their special evaluator process, the most frequently cited responses concerned their ability to effectively evaluate special educators. This finding is significant when considered with principals' belief that special educator assessment should differ from general educator assessment. Principals' belief that their current evaluation process works well for evaluating special educators is seemingly at odds with the need for differentiated evaluation of special educators. When asked to comment on weaknesses of the special educator evaluation process, principals noted the narrow focus of the current process. Criticizing the "one size fits all rubric,"

respondents highlighted the lack of alignment between their evaluation tools and student-specific disabilities (there may be no requirement to even specify the nature of a student's disability). One principal noted that, for example, if a teacher is teaching a student who is nonverbal, it is difficult to evaluate this teacher for student discourse, which affects the teacher's ranking in the distinguished category. Collaboration and communication were other important aspects of successful evaluation, as well as equity and fairness for students and teachers alike.

Findings also showed that principals ($N = 42$) who responded to this question believed that educator evaluation should be altered to take into account the specific roles and expectations of special educators versus general education teachers. When asked if the teacher evaluation process should be so altered, 42% ($N = 40$) of participants selected "agree," while 26% ($N = 25$) selected "strongly agree" to this question. In the comments section, study participants noted that "There is so much more variation in a special educator's teaching practice (coteaching, multiple grades/ages serviced in the same class, etc.)" (Participant 93) and because "there is much more nuance that goes into planning for special education teachers," these factors should be considered when determining the efficacy of a special education teacher's performance (Participant 60).

Participants agreed that grade level state assessments are misaligned with special needs students' abilities ($N = 31$ stated they "strongly agree), and that assessment of teacher performance are most meaningful and helpful when they are constructed on well-defined teaching standards ($N = 39$ selected "agree"). However, most participants disagreed that special needs students should be expected to achieve the same general-

topic area education goals as his or her counterparts, and that if these students do not reach their milestones, the teaching of these functional skills should be viewed as "wasted time." Many participants agreed that 'general' education students and special needs students must not be held to the same standards ($N = 33$). Likewise, participants agreed that special educators should not be held to the same standards as general educators.

In order to address the issue of unfair evaluation methods on special education teachers, the study's participants noted that teachers should be required to provide portfolios and students' weekly assessments, and be evaluated based on these scores, not based on traditional yearly tests and evaluations of students' scores. Furthermore, many participants agreed ($N = 32$) that classroom observations must provide the strongest source of evidence for gauging teacher performance. However, many participants noted that while observations are "essential," other factors such as "assessments, teacher attendance, teacher preparation, and teacher professional development" must be considered.

Interestingly, the majority of participants who responded to this question ($N = 51$) agreed that they find inadequate time to prepare for teacher evaluations, regardless of whether those evaluations are for special or general educators. Many of the participants ($N = 24$ who selected "agree"; $N = 21$ who selected "strongly agree") who responded to this question noted that their schools have inadequate school budgets and resources needed in order to provide effective and well-prepared teacher evaluations.

Descriptive Statistics Summarizing Research Questions

This section summarizes descriptive statistics findings for research questions. Upon completion of data collection, the responses which were derived from the survey instrumentation were analyzed with the Windows software package SPSS. The 4-point Likert scale was collapsed into two categories of agreement and disagreement. The following discussion presents descriptive statistics for each research question.

Research Question 1

How do elementary school principals perceive special education teacher evaluation, on basis of their effectiveness? The following responses indicate that principals believed SPED teachers should be evaluated using a different process than general education teachers (69.8% strongly agreed). They believed this process should reflect the unique challenges that diverse learners present. They also believed that in addition to good instructional practice being required, there should be alignment between grade level state assessments and capabilities of special needs students. Principals indicated that there may be misalignment of the state assessments and special needs students' abilities (89.9% agreed). Also, slightly more than half the principals (55%) disagreed that special needs students should be prescribed the same education goals as general education students for a given grade level (see Table 2).

Table 2

Responses Related to RQ1

Survey Question	Agree or Strongly Agree (%)	Disagree or Strongly Disagree (%)
Q2. Teacher evaluations must differ based on the differential roles and responsibilities of special educators (*N* = 63).	69.8	30.0
Q3. Educator assessment lacks precision when the evaluator does not have knowledge of specific practices which contribute to improvements in student performance (*N* = 62).	80.7	19.4
Q4. Teaching special needs students requires more than just good instructional practice (*N* = 62).	80.4	22.6
Q6. Grade level state assessments may be misaligned with special needs students' abilities (*N* = 59).	89.8	10.16
Q7. Special needs students must be expected to achieve the same education goals as 'general content' students at his/her grade level (*N* = 60).	45.0	55.0
Q9. The same student performance and student evaluation results should be used to assess teaching quality for both special and general educators (*N* = 61).	31.2	68.9
Q10. Teacher performance analysis is most meaningful and helpful when constructed on several rankings and well-defined teaching standards (*N* = 60).	91.67	8.33

Research Question 2

Which key indicators do elementary principals use to observe teaching quality?

Responses indicate that elementary principals perceive evaluation measures used to

assess teaching quality must be based on widely accepted standards. They

overwhelmingly believe that multiple measures of effectiveness must be used for

evaluation purposes (91.67%). While more than 90% believed additional measures of

teacher effectiveness should be used, including portfolios and students' weekly

assessments, three quarters of the respondents also felt that classroom observations

should be the primary source of information to measure teacher performance (75%).

Nearly two-thirds of principals disagreed that distinguished ranking awarded to teachers

was in large part due to grade level student growth (62%). These results in Table 3

summarize responses related to RQ2.

Table 3

Responses Related to RQ2

Survey Question Related to RQ2	Agree or Strongly Agree (%)	Disagree or Strongly Disagree (%)
Q11. Validated evaluation measures based on widely accepted standards are essential ($N = 60$).	90.0	10.0
Q12. A single test score cannot accurately represent teacher effectiveness ($N = 60$).	91.67	8.33
Q13. Teacher assessments must employ a range of measures of teacher efficacy, including portfolios and students' weekly assessments ($N = 60$).	91.67	8.34
Q14. Classroom observations are the only useful source of information for measuring educator performance ($N = 60$).	75.0	25.0
Q15. The distinguished ranking in tea cher evaluation is largely associated with grade level student growth ($N = 58$).	37.9	62.1
Q18. Teacher evaluation is objective when the evaluator has experience in teaching and assessing students with special needs ($N = 58$).	58.62	41.4

Research Question 3

Are the key indicators of educator quality different when principals evaluate general education teachers? Responses for the survey items summarized below in Table 4 show elementary principals' perceptions regarding the role of key indicators of SPED teaching quality. More than two-thirds of principals do not believe that all the influences of students' teachers, in addition to the one under evaluation, can be identified (70.9%).

Two-thirds of principals believe that special educators are evaluated as distinguished with

the same frequency as general education teachers (63.6%). These results are summarized

in Table 4.

Table 4

Responses Related to RQ 3

Survey Question	Agree or Strongly Agree (%)	Disagree or Strongly Disagree (%)
Q16. One may identify the influence of a special education students' other educators using teacher ratings ($N = 55$).	29.1	70.9
Q17. Special educators are evaluated as distinguished with the same frequency as general education teachers ($N = 55$).	63.6	36.4

Research Question 4

What barriers do elementary principals encounter in evaluating special educators?

The following responses summarized in Table 5 show principals' perceptions of

obstacles that impact their ability to assess special educator effectiveness. Less than one

principal in five believes he or she lacks the qualifications and experience to evaluate

SPED teachers (17.7%). They cited budget constraints and the demands of their position

as barriers instead (73.8%). Respondents were almost evenly divided as to whether time

pressures and inadequate assessment tools posed problems for them (51.7%).

These results are shown in Table 5.

Table 5

Responses Related to RQ4

Survey Question	Agree or Strongly Agree (%)	Disagree or Strongly Disagree (%)
Q20. I do not have the prerequisites (qualifications, experience etc.) to conduct special education teacher evaluations ($N = 62$).	17.7	82.3
Q21. There is a lack of employer support ($N = 58$).	39.7	60.3
Q22. There is inadequate time to prepare. ($N = 60$).	51.7	48.3
Q23. The assessment tools are inadequate ($N = 60$).	48.3	51.7
Q24. Inadequate school budget and resources ($N = 61$).	73.8	26.2
Q25. Government regulation and policy ($N = 49$).	57.3	42.7
Q26. High workload or workplace responsibility ($N = 60$).	81.7	18.3

The study explored associations between elementary principals' perceptions of special education teacher evaluations in reference to general education teacher evaluations and measurements of teacher quality, effectiveness, and barriers to effective assessments. In summary, the findings demonstrate the following:

- Elementary principals do perceive the evaluation process for special education teacher effectiveness to be the same as the evaluation process for general education teacher effectiveness.

- Elementary principals use evaluation measures that are not applicable to special education teacher assessment.

- Key indicators of teaching quality are different for the evaluation of general educators when compared to special educators.

- Elementary principals encounter barriers in their mission to evaluate special education teachers in an appropriate manner.

These findings aided in creating a position paper. The data are used to inform policy development and promote further improvements to evaluation systems. I am entering these findings in the position paper with the policy recommendations presented in Appendix A.

Conclusion

Section 2 described the methodology the study employed along with procedures for participant selection and recruitment. The section also discussed data analysis and data collection as well as study limitations. The following section 3 presents a description of the project.

Section 3: The Project

Introduction

In Section 3, I review the position paper's components, with the policy recommendation presented in Appendix A. The section includes a discussion of research related to exploring elementary principals' perceptions of the process for evaluating performance of special education teachers. The section describes the project, its goals, and outcomes. This review also includes a discussion of the rationale for choosing this project to address the problem of investigating elementary principals' perceptions of special educator evaluations. The review of literature expands upon themes introduced in Section 1, including a discussion of FFT used as a protocol for observation in the study site. This section also presents the project design and implementation plan including required elements, existing levels of support, potential obstacles, and likely solutions. Finally, this section discusses social change implications and describes this project's significance on a national and local level.

Project Description and Goals

This investigation's purpose was to gain insights into the assessment process for special educators by analyzing elementary principals' perceptions of existing procedures. Additionally, the purpose of this research was to create a position paper—that is, a paper or research study completed to generate support on an issue. The main purpose of a position paper is to describe an issue and explain the rational reasons for choosing a specific stance on the issue. This study's objective has been to increase understanding of

how principals balance competing requirements for effectiveness, fairness, and accuracy in measuring special education teacher performance.

Position papers use evidence and authoritative references to support a position to show why that belief is the best method of moving forward with that particular issue. For example, this position paper sought to increase general awareness and understanding about the failed methods for special teacher evaluation processes, and spur growth and development in improving special education teachers' evaluation systems due to an increased focus on the topic. This research particularly focused on finding the thoughts and assumptions of principals, who are most likely the ones rating special education teachers' performance. It was believed that elementary principals' survey responses would help to answer the study's research questions examining special education and general education assessment, key indicators of teaching quality that principals seek to observe, and barriers to effective evaluation. Major outcomes of the project recommendations are insights that would help the professional development of all teachers, particularly special educators, in their efforts to support student achievement.

Rationale

I decided to write the position paper as the best means of addressing the lack of research on how the teacher evaluation needs to be different to reflect the skills and practice of special education teachers. Even though nearly all 50 states have begun the process of implementing new evaluation systems to meet federal mandates, few appear to have developed assessments that differentiate between the roles of special and general educators (NCCTQ, 2010). By asking one school district's elementary principals to

discuss their perceptions of the evaluation process through survey responses to open- and closed-ended questions, this position paper on teacher evaluation helps to address concerns that performance evaluation criteria are not appropriate for reflecting the responsibilities of special education teachers, which diverge considerably from general educators. This research project expands what is known about evaluation systems for public educators and helps address the challenges of accurately linking student growth to teacher effects. The position paper, which is a key product of the research process, presents stakeholders with information that can be used to inform policy development and promote further enhancements to evaluation systems.

Review of the Literature

A review of extant scholarship summarized challenges associated with teacher evaluation systems, and teacher and administrator perceptions of those challenges. The project primarily focused on peer-reviewed sources that were less than five years old. Because this project conducts a review of policies dating back two or more decades, the literature review also refers to those legislative mandates as well as position papers from think tanks and policy analysts having requisite subject expertise. This review focuses on the specific genre of this project: a position paper with policy recommendation. To create an effective position paper with policy recommendation, a literature review was conducted. Online searches used keywords and phrases such as *policy recommendation, policy development, policy changes*, and *policy framework*. All online searches were conducted through Walden University Online Library and Google Scholar.

Teachers' Perceptions of Current Evaluation Systems

Teacher perceptions of the use or effectiveness of evaluation practices are not particularly positive. In a survey supported by the DoE of 10 Arizona school districts, it was found that 32 percent of teacher survey respondents did not believe the performance classification they received accurately reflected their overall performance after their schools had put a new multiple-measure evaluation method in place, in 2012/2013 (Ruffini, Makkonen, Tejwani, & Diaz, 2014, p. 1). Further, only 39% agreed that their evaluation was accurate, while 30 percent were undecided (Ruffini et al., 2014). Teachers in five of the 10 school districts voiced concerns over consistency in classroom ratings by principals, while teachers in 3 districts were concerned about the type and amount of observations which were necessary in order to rate teachers' performance in an accurate manner. Both educators and principals who were focus group participants in the same study voiced concern over the lack of calibration among evaluators. Teachers from two districts in the focus groups noted that principals needed additional training to evaluate teachers consistently. Among teacher survey respondents, only 51% responded that the amount of formal observations they had were sufficient to assess their performance (p. 6), while 26 percent disagreed. Teachers from three districts felt that the number of high-need students they taught precluded standardized test score improvement from being their highest priority as educators. When asked if they agreed that the newer teacher evaluation process constituted improvement over prior methods of evaluation, 30% disagreed, while only 25% agreed. When asked if the new teacher evaluation process was fair, 31%

disagreed, while 34 agreed (p. B-3) (Ruffini et al., 2014). Harris et al. (2014) uncovered similar findings.

Teachers also expressed doubts about principals' ability to evaluate educators across multiple grade levels or subjects (Ruffini et al., 2014). Survey data in Tennessee showed similar negative perceptions. Results from a 2012-2013 survey showed that 50% of respondents disagreed with the statement that they were satisfied with the evaluation process used in their school (Schwartz, 2013). When asked if the processes used to conduct their own teacher evaluation was fair, only 34% agreed with the statement (Schwarz, 2013). Nor do some principals themselves give the new teacher evaluation system high marks. In an examination of the Ohio teacher evaluation system, Kowalski and Dolph (2015) found both teacher and principal attitudes and feelings towards evaluations have been more negative than positive.

Teachers also weighed in on the debate over the effectiveness of teacher evaluations through their unions. Teachers' unions have voiced concerns over the fact that some teachers have many more students with special needs or challenging home circumstances than others (Harris et al., 2014). Unions also noted the unfairness of judging teachers by the scores of students they do not even teach, as required by some states' evaluation systems. Such concerns have resulted in more than a dozen lawsuits over new evaluation systems (Sawchuk, 2015) including those filed in Tennessee and Florida in 2014 and 2013 respectively (Sawchuk, 2014).

Maharaj (2014) called for considerable reform to educator evaluation in his discussions of evaluations from the administrator's perspective. His findings drew from

research studies, including the Canadian province of Toronto, which documented that

evaluations were perceived to be time-consuming and less than useful. Teachers felt the

evaluations were neither objective nor accurate. This ineffective public policy

combination gives the appearance of accountability by political and education leaders

without producing actual benefits for teacher practice (Harris et al., 2014). New

evaluation systems are criticized by some as being too lenient or incompletely

implemented, while others criticize them as being unfair or counterproductive (Sawchuk,

2014). Given the abundance of such widely varying views, the evaluation report which

this study produced is appropriate for analyzing this contentious topic.

Evaluating Special Educators

With all the controversy surrounding teacher evaluations in general, it is not

surprising that special educators in particular question the efficacy of current evaluation

systems (Harris et al., 2014). Measuring teaching effectiveness in special education

presents unique challenges because of special education teachers collaborating to share

the load of creating instructions, handling the various special education student abilities

and levels, and because of teaching special education students happens over a number of

settings (such as independent classrooms, resource classrooms, and in classrooms

dedicated to coteaching) (Jones & Brownell, 2014). Additionally, academic

accomplishment is one of many desired outcomes for students with disabilities, who may

also have goals that include positive communicative, behavioral, adaptive, transition

social results (Jones & Brownell, 2014). It is therefore possible that special educators

cannot demonstrate all of their responsibilities, thereby calling into question the appropriateness of the FFT's role in evaluating special educators.

Even with a greater number of inclusive options and more disabled students having access to grade level standard curriculum, achievement gaps continue for special education students. The National Assessment of Educational Progress for 2009 reported a variance of 35 points for general education students' reading scores compared with those of special education students (Sledge & Pazey, 2013). Math score differences revealed 21 points of achievement gap among fourth grade students, which increased to 58 points in grade eight. These differences in student achievement underscore the challenges involved in designing systems to measure student growth as promoted by special education teachers (Darling-Hammond, 2015; Goe et al., 2014).

Minnesota acknowledged the unique challenges that evaluation poses for special education teachers in a recommendation by a state task force of special education teachers and leaders. A key finding of this group held that special educators must be involved in developing, implementing, and assessing the any evaluation to which they may be subjected (MDE, 2014; Spina, Buckley, & Puchner, 2014). The group further recommended that a "qualified and trained summative evaluator such as a school administrator" should develop an awareness of various roles performed by special education teachers (Spina et al., 2014). The group also called for measures of student growth to be fair and accurate and reflect the special educators' specific contribution to such growth. A focus group consisting of teachers and administrators in a school district

in southwestern Illinois also expressed concern about evaluating special educators in a manner unaltered from those employed to evaluate general educators (Spina et al., 2014).

Classroom observations of students with disabilities (SWDs) are affected by problems with both validity (teachers' evaluations presenting a potentially inaccurate, or incomplete assessment of classroom education) and equality (causing negative incentivization to address SWDs' needs) (Harris et al., 2014; Buzick, Jones & Turkan, 2013; Ravitch, 2016). Studies highlight the difficulty of teaching observers to score in reliable ways (Bell, Gitomer, Hamre, McCaffrey, Pianta, & Qi, 2012) and to reduce variation in the scores for teachers so observed (Charalambous, Hill & Kraft, 2012).

The Principal's Role in Special Educator Evaluations

As mentioned above, principals are generally the only evaluators in charge of assessing a special education teacher's performance. Derrington (2014) examined principals' perceptions of teacher evaluation policy implementation. Three of the four districts Derrington studied added personnel specifically to evaluate teachers, thereby reducing the time principals spent evaluating. Murray (2014) also found that one of principals' many responsibilities is to evaluate special education teacher performance, even though they may not be qualified. Requiring evaluators, especially principals, to follow a set guideline or undergo additional training increased rater reliability and improved the quality of special education teacher evaluations (Derrington, 2014; Murray, 2014).

Getting teachers, principals, and school administrators at Midwest School District (MSD) on the same page, so to speak, would allow the school district to make massive

bounds in improving the quality of its special educator evaluation methods/system. Specifically, rater reliability is critical for fair special education teacher evaluation, and increasing rater reliability and competency is something that can only be achieved through a policy change. This subject is therefore paramount to address in the policy recommendation for MSD.

Alternatives for Policy Recommendation

Jones and Brownell (2014) stated that in-classroom observation for evaluating special educators is a policy that is in dire need of change. However, when presenting a policy recommendation, it is necessary to create alternatives for policy recommendation in case the first policy is not received well, cannot work within the school district's framework, or is ineffective or incompetent for any reason (Bardach & Patashnik, 2015; Firestone, 2014). Additional evidence and research on a topic must be presented in more than one way in order to help the target audience make an informed decision (Bardach & Patashnik, 2015).

Project Description

Potential Resources and Existing Supports

The researcher can use the Midwest Teachers Union (MTU) to provide contact information for principals to participate in follow-up research. The MTU can also provide additional resources and assessment-geared information, as well as that focused upon the evaluation of both general and special education teachers. The researcher can also obtain information on enhanced evaluation practices from colleagues and other administrators. Educators and administrators will be solicited directly, and if necessary, media attention

will be brought to bear on the school district. As has been considered, such attention may aid in bringing to light the range of deficiencies in evaluation of special educators considered in this work, as a means of exerting change pressures on the school district under consideration.

Potential Barriers

Barriers include determining which survey respondents have already identified enhancements to the evaluation process. Because the Survey Monkey platform protects the confidentiality of respondents, it may not be possible to directly contact survey participants who are already using improved assessment techniques. Another barrier to research is resource constraints. Case studies, focus groups, and interviews are all labor-intensive undertakings that would require the use of additional personnel to assist in data collection. Coordinating with administrators to schedule more research is another hurdle.

Proposal for Implementation and Timetable

The Implementation and Timetable would depend on the project design selected for the next phase. A single case study consisting of an observation of one administrator and one special education teacher could require as little as one month to complete. Focus groups or interviews would require several months to arrange, collect and analyze data, and summarize the results. Following the collection of such data, it will be implemented toward the exertion of change pressures toward policy reform – including toward superior special educator evaluator education and expertise – at the school district under consideration. The use of discrete data collected within this school district will aid

immeasurably in this effort. However, because the policy recommendation implementation is a necessarily political effort, no clear timetable can be anticipated.

After investigating the problem, the next steps for the study involved determining specific practices that will result in enhanced and improved evaluations for special educators. In their discussion of the strengths of their current evaluation process, principals referred to a number of factors that accounted for their success. These factors included:

- Including administrators as part of the evaluation team who have backgrounds as special educators

- Using measures other than test scores to evaluate SPED teachers

- Including a SPED (Special Education) addendum to the Recognizing Educators Advancing the City (REACH) process, the recently redesigned MSD teacher evaluation system

- Adapting the Danielson Framework rubric descriptors for SPED teachers

- Preparation and training for administrators to perform observations

Administrators identified evaluation practices they believe enhance their ability to identify differentiated tasks that special educators perform. Which specific measures do they use in addition to test scores? Which adaptations to the Danielson rubric do they use? How did they modify the SPED addendum for REACH? What training have they implemented to better prepare administrators for collection of observation evidence and rubric use? These questions remain to be answered.

Gathering this additional information required contacting principals and performing more research that provides details of improved evaluation practices. Additional research should consist of case studies, focus groups, or interviews to provide rich details that can help other principals achieve improved assessments as well. The policy recommendations that are outlined demand a close coordination with the local school district, conducted as a matter of *political advocacy*. In essence, it is anticipated that the district (and the elementary schools it contains) will not inevitably be receptive to engage in the change recommendations which follow from this research. Change can be difficult for any bureaucracy, even when imposed from within (or by government mandate), meaning that any change effort which results from external advocacy must be presented in as straightforward and forceful a manner as possible, if it is to be effective. This will demand direct lobbying for these change measures, but if necessary, local media may be solicited to aid in these efforts. Special educator evaluation is a concept which might be made easily-digestible by the public, and the deficiencies therein may be properly-framed in the news media as a manner of child welfare, especially as it pertains to the welfare of students with learning or developmental disabilities. Though such efforts may be unnecessary, newspaper or local television reporters may be engaged to bring local attention to the deficiencies in evaluative capacity held by local principals.

Responsibilities and Roles Held by Students, and Others

The researcher must identify and contact additional study participants. The researcher is also responsible for engaging assistants to help with data collection and analysis. Study participants would be responsible for reviewing transcripts and providing

feedback. In the course of the advocacy which will follow from the results of the data-collection process, the responsibility of this student is similar to that of other political actors seeking to change entrenched bureaucracies: In essence, the researcher is free to act within the boundaries of the law, provided that such action is conducted in an ethical manner. To this end, solicitation of local media support will be appropriate.

Project Evaluation Plan

As noted above, the purpose of this research was to create a position paper to describe an issue and explain the rational reasons for choosing a specific stance on the issue. This project's deliverable was selected to be a position paper with a policy recommendation for MSD, specifically to create a new proactive approach to improving methods and systems of special educator evaluation. This research particularly focused on finding the thoughts and assumptions of principals, who are most likely the ones rating special education teachers' performance. Position papers use evidence and references to support a position to show why that belief is the best method of addressing that issue.

Position Paper

This position paper sought to increase general awareness and understanding about the failed methods for special teacher evaluation processes. It is hoped that by increasing awareness of this topic, further research would be completed, encouraging growth and development in improving special education teachers' evaluation systems. As is discussed below and in the review of the literature section in section 1, many research

studies have indicated that a problem exists regarding a paucity of applicable and relevant evaluative systems for special educators.

To address this issue, this study focused on evaluating and analyzing the thoughts and assumptions of principals, who are most likely the ones rating special education teachers' performance. It was believed that elementary principals' survey responses would help to answer the study's research questions examining special education and general education assessment, key indicators of teaching quality that principals seek to observe, and barriers to effective evaluation. The position paper's final goal is to provide major recommendations and insights that would help the professional development of all teachers, particularly special educators, in their efforts to support student achievement. Before it is possible to create a position paper, however, it is necessary to understand the background of the problem.

Recent trends toward teacher accountability that began in the 1990s (Murphy & Hallinger, 2013) have resulted in school districts needing to implement the most effective possible means of evaluating teachers (Evans, Wills, & Moretti, 2015). There is general consensus that the U.S. system of teacher evaluation is ineffective (Murphy & Hallinger, 2013; Kane & Staiger, 2012; Sawchuk, 2015). Prior to 2008, the traditional teacher evaluation system that MSD used identified nearly all (93%) of their educators as superior or excellent, despite the fact that two in three (66%) of MSD schools failed to meet standards set at the state level (Sartain et al., 2011). The movement calling for effective teacher performance evaluation took on added urgency with passage of policy initiatives like RTT and NCLB (Callahan & Sadeghi, 2015; Polikoff & Porter, 2014).

Across the country school districts, states, and government agencies noted the absence of fair, reliable, and effective means of assessing teacher performance. In fact, from 2009 to 2015, states which mandated that student achievement levels be taken into account in teacher evaluations increased from 15 to 43 states (Jacobs, 2015). A number of researchers have challenged the validity of value-added data (Berliner, 2013; Kersting, Mei-kuang & Stigler, 2013). Moreover, what constitutes effective measurement in teacher evaluation has yet to be decided (Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012; Garrett & Steinberg, 2014). Because of these shortcomings, there is a need for an evaluation study, a genre well suited to investigating this problem.

Policy Recommendation

According to research, a policy recommendation should be created in sections, guaranteeing that every aspect of presenting a policy recommendation is included in the recommendation (Bardach & Patashnik, 2015; Firestone, 2014). The overall goal of a policy recommendation is to encourage key players (in this case, administrative personnel at Midwest School District) to use data from this study and other studies to identify better methods to evaluate special education teacher performance (Firestone, 2014). A well thought out, complete policy recommendation includes several sections: (1) define the objective, (2) target an audience, (3) clearly present the issue, (4) provide alternatives, (5) provide cost effectiveness, (6) works with other strategies, (7) provide similar examples, (8) written in simple language, (9) support social change, and (1) emphasize taking action (Bardach & Patashnik, 2015). Creating an effective policy recommendation, however,

can only be possible while understanding the views of current modes of evaluation held by educators themselves.

Project Implications

Further research will be required to determine what works and what does not work to improve the special educator evaluation process. This research will require not just the administrators' perceptions, but those of special education teachers and other key stakeholders, including families of special needs students. This goal-based exploration needs to examine whether the newly implemented evaluation processes align with the findings of the original project. One of the challenges with conducting additional research will be to establish relevant and specific performance goals to evaluate, as well as to determine what measurements to use. In addition to considering principals' perceptions, an ongoing literature review is necessary to monitor the outcomes of additional studies as more and more school districts experiment with different assessment models.

Social Change Implications

In the larger societal context, it is important that everyone in our society be given the opportunity to reach their fullest potential. In the U.S., education is guaranteed by state, federal, and local mandates. But even though these rights are recognized by policymakers and stakeholders, it is not clear that special educator evaluation practices advance this agenda. One of the most significant ways of promoting the success of special needs students is to develop instructional practices that accurately and fairly evaluate SPED teachers. The proposed change will begin as a local effort, one which is not guaranteed to be implemented, even with the change pressure assistance of the local

community or local news media. However, if these efforts are brought to bear, they may result in changes not just to local educational evaluation policy, but they may be used as grounds by which national educational policy may be changed as well. Since the inception of performance-based evaluations in American education, many critics have decried their essential deficiencies. However, true change is not a matter of simply criticizing an existing system, but often a matter of presenting a favorable alternative. If national educational policy is to be changed, this local pilot educational evaluation reform initiative may well prove a strong body of evidence in its favor.

Far Reaching Impact

This study contributed to gaining insight into the evaluation process for special educators. There is widespread agreement that the special educator's contribution to student growth is difficult to identify and quantify using existing tools and measurement. The current process used by most respondents does not account for variances in diverse learner's abilities, the difficulty of special needs students achieving grade-level performance, effectively measuring student growth, particularly whole child development, or the effects of coteaching. The current process also does not lend itself to professional development for SPED teachers, providing infrequent opportunities for meaningful participation, discussion, and feedback. Improving the evaluation process benefits students, teachers, families, and communities benefitting by enhanced education practices that promote student and teacher growth. Policy-makers will also benefit from an ability to demonstrate accountability to taxpayers and community partners.

Community Impact

Provided the local school district shows responsiveness to reform initiatives as proposed, by which principal-evaluators receive greater education and expertise to aid in their evaluation of special education students, vast benefits stand to result. That said, this is not necessarily a change which will come immediately, or without challenges from the school district bureaucracies who may interpret this – political – effort as a challenge to their power or to a preferred *status quo*. In this light, the necessity of soliciting assistance from the local community or a spotlight from the local news media, may be necessary. Efforts such as these often demand an entire community coming together in order to identify a major problem, and insisting that those with the power to correct current deficient policies do so. In this way, the impact of these policy recommendations on the local community may serve to engender a stronger connection, and state of awareness, between local parents and their children's schools – and educators – especially if they are parents of children with disabilities.

Importance of the Project to Stakeholders

Providing a policy recommendation for an issue is commonly the main goal of a position paper (Firestone, 2014). Therefore, the policy recommendation presenter must be aware of the key stakeholders and how the information will reach the targeted audience (Bardach & Patashnik, 2015). For this study, the target audience is school administrators for Midwest School District, as mentioned above. This group has been targeted as the key audience for this position paper's policy recommendation because, as

they are in charge of the school district, they are most competent and able to make changes in the Midwest School District.

Understanding where an organization's power lies is paramount to finding the key audience (Bardach & Patashnik, 2015; Firestone, 2014). Once the main stakeholders and audience is defined, it is important to take necessary steps to present the policy recommendation in such a way that the message is clear, easy to understand and the audience members are aware of its importance (Bardach & Patashnik, 2015; Firestone, 2014). As was mentioned above, a competent policy recommendation includes ten key points including defining the objective, clearly presenting the issue, providing examples, and writing the policy in simple language, alongside emphasizing taking action (Bardack & Patashnik, 2015). It is critical to convey the importance of this policy recommendation in simple, concise language in order for the recommendation to be effective to the target audience (Bardach & Patashnik, 2015; Firestone, 2014).

Conclusion

Section 3 presented a description of the project and its finding. The section summarized results of responses to closed-end survey questions by 97 administrators. The section also presented an analysis of themes revealed in responses to 5 open-ended survey items. Section 3 also discussed tasks involved in implementing the next steps of this research project. Section 4 will reflect on results and themes, as well as conclude the study. The next section will also address its limitations, project strengths, and recommendations to address the issue of special educator evaluation.

Section 4: Reflections and Conclusions

Introduction

In this section I discuss and reflect upon conclusions derived from this study, through assessing both this study's purpose, reviewing its research questions, and providing a summary of its methodology and key findings. Finally, implications for social change and applications to the educational field are discussed. This study's purpose was to develop insights into the evaluation process for special educators by analyzing elementary principals' perceptions of existing procedures. The goal was to further understand how principals balance competing demands for effectiveness, fairness, and accuracy in measuring special educator performance. It was believed that principals' survey responses would help to answer the study's research questions examining special education and general education assessment, key indicators of teaching quality that principals seek to observe, and barriers to effective evaluation.

I collected qualitative data from the participants to develop a rich dataset that would allow the researcher to explore principals' perceptions about special educator evaluation. Open-ended survey questions asked participating principals about strengths and weaknesses of their special educator evaluation process, desired changes, and comments on evaluation practices and related issues. Closed-ended survey questions asked principals to rank, using a rating scale, their agreement or disagreement with applicability of evaluation principles to special educators, indicators of teaching quality, and barriers to effective evaluation. Qualitative analysis revealed themes principals considered important:

- Differentiated evaluations to reflect the nature of the special educator's role

- The need for inclusion and accommodation for diverse learners

- Supporting student growth in academic and social dimensions

- Employing measurements that were effective, fair, accurate, and consistent

- Validating good instructional practice

- Accountability for student learning outcomes

- The impact of resource constraints

- The impact of the administrators' qualifications on the special educator evaluation process

- Barriers to effective evaluation

Descriptive analysis of responses to closed-ended survey items appeared to confirm the concepts which were laid out in the research questions.

Project Strengths and Limitations

The high points of the policy paper written from the findings of the research provide strong evidence to support the recommendations submitted in the project study. The strength of the position paper is derived from its capacity to provide insights that could lead to resolving pressing issues American education, particularly that of students with special needs. The educational field faces considerable problems that range from facilitating accountability to retaining qualified special educators to inadequate resources needed to promote student success. The principal is at the nexus of all these demands and

must successfully resolve competing requirements that teacher evaluations be effective, fair, and accurate. Exploring elementary principals' perceptions of how they meet these challenges contributes to developing effective education practices.

Legislation mandating that teacher effectiveness be aligned with student success has generated a number of approaches to teacher evaluations. New tools and processes are only beginning to reach the stage where there is sufficient data from multiple school years to allow for meaningful analysis. The present study comes at a time when additional insights into the efficacy of the assessment process can prove useful.

The debate around teacher evaluations involves controversy and opposing approaches. Indeed, recent developments in teacher assessment are still unfolding as more data becomes available concerning tools and practices currently in use. Administrators, special and general educators are all motivated to improve current processes, and this study provided a vehicle to advance that improvement.

Elementary principals are uniquely positioned to add to the discussion around special educator evaluation. They bring an important perspective concerning which tools and which procedures work, as well as what does not work. They are able to discuss lessons learned from early implementations of tools such as Danielson's addendum for special educator assessment. Similarly, they continue to experience the ongoing need to improve tools and processes currently in use.

This research was limited in several important ways. While the project considered elementary principals' views regarding educator evaluation, and its efficacy, the project did not address the similarly important perspective of educators themselves and how they

perceive evaluation practices. Another important group of stakeholders, parents, also had no input reflecting their concerns about how the evaluation process best serves the needs of their children. Effective education practices ideally should include all these viewpoints.

Another limitation resulted from the approach which limited the project to primarily closed-ended questions. Survey items, by definition, address a limited number of topics and offer a limited range of possibilities for expressing agreement or disagreement. This approach may not address all factors worth considering for further exploration. This limitation was only partially addressed by including open-ended questions as part of the survey data.

An additional limitation of using surveys was the inability to probe participants' comments and elicit further details from them. Participants' responses indicated that the survey had only scratched the surface of critical topics, and they would have welcomed an opportunity for more in-depth exploration. Access to participants was another research limitation. Contacting principals involved working through the school district for approval to implement data-collection. In addition, this study was mounted during the summer break, which possibly limited the response rate. Without these limitations it may have been possible to collect more responses from a wider audience of principals. Increasing the response rate might also have lessened the possibility of self-selection bias. Another limitation arose with the use of one urban school district in one state, as well as the small response rate within that school district. These limitations negatively affected the generalized ability of the project findings.

Recommendations for Alternative Approach

Many limitations could be addressed by using a different research design. For example, a case-study approach could be used to focus on teacher evaluations at one school. This approach could allow the researcher to observe classroom observations, evaluation tools, and modifications for special needs students during the evaluation of multiple educators. This approach would generate more in-depth and varied data for consideration.

Interviewing or employing focus testing on principals and teachers would also aid in reducing limitations identified in this data-collection process. Interviews would be helpful in that they might encourage subjects to provide information in a more informal manner, which might lead to greater data fidelity, as by reducing group-based pressures to accede to demands of political sensitivity. Principals' lack of special educator experience, for instance, might be a topic better-elucidated through confidentiality.

Longitudinal studies over time could also be used to overcome survey limitations. Studies over a period of several academic years could show the development of proficiency with assessment tools, the impact of accumulating more data for inclusion in assessing teacher effects, or an increase or decrease in satisfaction with the existing process.

The study might be undertaken to explore teachers' perceptions of the evaluation process, perhaps focusing on topics such as coteaching or professional development. An additional alternative approach would involve exploring parents' perspectives on teacher evaluations, perhaps focusing on the IEP process or nonclassroom contributions to

student growth. Any of these redefinitions might be investigated using a different research methodology as previously discussed.

Scholarship

I learned valuable lessons about scholarship as a result of this study. One of the most significant lessons was an appreciation for viewing information from different perspectives to arrive at different conclusions. An extensive review of the literature revealed how differently people can perceive the same issues. Survey responses showed similar variation and helped me develop an appreciation for the process.

Exposure to an array of arguments while conducting this study prompted me to approach all research with healthy skepticism. I learned to appreciate both different and similar priorities that administrators and educators bring to a discussion of teacher evaluations. I also learned to incorporate as many voices as possible into the debate surrounding teacher evaluation. The development of policy recommendations taught me the importance of considering a wide body of stakeholders, and the significance of framing a given issue in as persuasive a manner as possible. Because change (especially in an 'entrenched' government bureaucracy, as this school district) will not automatically occur due to singular external advocacy, this process allowed me to discover the importance of outlining a particular grievance, and of framing the problem to be solved – and its solution – in the clearest possible terms. Such 'framing' can be used as a 'vehicle' by which the most possible allies can be solicited to aid in the necessary change effort.

Project Development

This project taught me how to expand an idea into a set of research questions. I began this project with a belief that teacher assessment tools and techniques needed improvement, but this belief was not a scholarly argument, nor much more than a break room conversation. However, this study helped me learn to frame my arguments in a way that allowed for their confirmation or rejection. This project taught me how to systematically and methodically review scholarly literature and consider competing and opposing ideas. This project also brought me the satisfaction that comes with seeing a project through to completion. Not least of all, this study made me more comfortable with making challenging decisions about handling objections and overcoming obstacles and resource constraints.

Leadership and Change

This project reinforced my belief that championing change and showing leadership requires boldness and determination. I have always been drawn to the teaching profession by a conviction that student needs come first, and I was comfortable advocating for change on their behalf. Nevertheless, testing, accountability, policy and procedures, budgets, and resources all have an impact on student success. Effective leadership involves the ability to manage all these realities and to do what is required even under challenging circumstances. I was fortunate enough to embrace opportunities to learn more about meeting adequate yearly progress. This project highlighted for me the constant tension between accountability for student performance and other important goals of fairness and accuracy in teacher assessment. The study expanded my

appreciation of the leadership skills and my willingness to embrace change that the educational field imposes on principals and policymakers.

Reflection on Importance of the Work

Analysis of The Self as a Scholar

As an educator and scholar, I learned how important continual learning can be. I found the value of suspending my own opinions to allow me to better appreciate the opinions of others. Even though I brought my own ideas about teacher assessment to this project, I learned to bracket my own feelings and to set them aside to more thoroughly investigate other points of view. I learned the value of keeping an open mind.

Analysis of The Self as Practitioner

This study reminded me again of why I was drawn to the field of educating special needs students. There are challenges to be sure, but the satisfaction of helping students reach their fullest potential more than makes up for the demands. Further, in the spirit of dedication to life-long learning, I appreciated opportunities this project gave me to learn how other educators solved problems associated with teacher evaluations.

Analysis of The Self as a Project Developer

This process allowed me to broaden and sharpen my skill set. I found myself using creative ways to access people within the school district and teacher's union who could further the goals of this research. I took advantage of my network of professional contacts and my colleagues for brainstorming and to find support and tools for this project. I also learned how to reframe findings of a study to interpret inconclusive results and analyze what could be done differently in research projects in the future.

Implications, Applications and Directions for Future Research

This study's outcomes shed light on answers given by elementary principals to

research questions under investigation:

- Principals do not perceive the evaluation process for special education teacher

  effectiveness should be the same as the evaluation process for general education

  teacher effectiveness.

- Principals believe they use evaluation measures that are not necessarily applicable

  to special education teacher assessment.

- Key indicators of teaching quality are different when evaluating special educators,

  as differentiated from general educators.

- Principals encounter barriers to appropriately evaluate special educators.

    Many of the study's participants felt that the special educator's unique role in

promoting the growth of special needs students did not accurately account for using their

current tools and indicators of teaching quality. This perception suggests a need for

further refinement of assessment approaches.

    Future research can inform such refinements by bringing in additional

perspectives on teacher evaluation. There is a need to explore how teachers feel about the

evaluation process. Educational improvement will not occur without significant buy-in

from teachers regarding the effectiveness of evaluation systems. Parents should also be

brought into the conversation that shapes how the special needs of their children will be

accommodated in a way that promotes inclusion and accountability. Parents need

reassurance that the best interests of their children are at the center of any discussion around teacher assessment.

Finally, no conversation about improving education practices is complete without a discussion of barriers to effective teacher evaluation. Whether barriers take the form of administrators who lack a background in special education or who face time and budget constraints, it is important to acknowledge the existence of barriers in order to develop the means to address them.

Potential Impact for Positive Social Change

This work has explored how elementary principals perceive the process of evaluating special educators that, in turn, contributes to promoting student growth and improving outcomes for students, their families, the community, educational practitioners, and stakeholders. The study accomplishes this by expanding discussion within the educational field regarding assessment approaches. More accurate measurements of student growth are necessary to help students achieve their fullest potential and to assist in the professional development of educators. Moreover, improved evaluation tools and processes promote accountability so that exceptional teachers are recognized and teachers who need to improve will understand what is expected of them. Teachers who believe they are evaluated in ways that are fair and accurate are more likely to remain as special educators, helping to minimize the existing shortage in this specialization (Harris et al., 2014). Finally, policymakers have additional data to inform their decision-making**.**

Methodological and Theoretical Implications

The project employed a research approach driven by determining the nature of the

problem, and as such, a qualitative methodology was necessary. Likewise, many other

articles examining the relationship between perceptions of special education teacher

performance and reviews (i.e., Holdheide & Reschly, 2010; Jones & Brownell, 2014;

Schulze, 2014) utilized qualitative methodology and theory-driven evaluation practice

(Coryn et al., 2011). It is by using these specific tried-and-true methodologies and

theoretical practices that they confirm or falsify the research problem and phenomena.

In the scope of the current project, the qualitative methodology focused on the quality,

not quantity, of the participants' (principals'/evaluators') responses. This required deep

in-depth questions of the participants, which allowed the researcher to codify their

responses based on theme or concept, which contributed to a deeper understanding of the

problem and their (the participants') perception of it. To put it another way,

understanding the logic of these educational practices requires an in-depth understanding

and close analysis of the field and how the special education teacher evaluation process

works. Therefore, through research and application of the findings it becomes necessary

to understand principals' perceptions of the roles, responsibilities, and accomplishments

of special education teachers, as well as how external environments (i.e., the classroom

environment, special education students' disabilities, etc.) will affect the special

education teacher's efficacy.

In understanding these key factors, the implications are that by understanding

social practice and evaluator comprehension, the closer researchers get to finding and

applying solutions to these real-world situations. Previous studies failed to effectively evaluate educators of students with developmental or learning disabilities. The implications of this study are, therefore, that by fostering evaluators to learn about the roles of a special education teacher, they may then evaluate and judge special education teachers more fairly and adequately.

The theory that evaluators (principals) often lack the expertise in special education necessary to mount an effective evaluation served as a framework to understand what changes need to be made to the current systems of evaluating special education teacher performance. Because the study hinged on this understanding, it stands to reason that change in this arena will not be possible without thorough analysis and understanding on the part of the evaluators. Furthermore, as this theory was explored and found effective in works by Jones and Brownell (2014) and Schulze (2014), it is reasonable to expect similar findings in this study.

Conclusion

This section includes a summary of reflections and conclusions resulting from the findings of the project. The guiding questions and study purpose were reviewed. This section also presented a summary of the methodology and key findings. Finally, implications for social change and applications to the educational field were presented.

The purpose of this study was to capture elementary principals' views, investigate how special educators' evaluation can be improved. From this point, progress can be achieved by incorporating the perspectives principals offered on effective teacher assessment and building upon this information. In conclusion, when asked the strengths

of their special evaluation process, an elementary principal from MSD interviewed for

this study stated: "They believe in our students."

References

American Statistical Association. (2014). ASA statement on using value-added models

for educational assessment. Retrieved from

https://www.amstat.org/policy/pdfs/ASA_VAM_Statement.pdf

Anderson, K., Smith, J. D., Olsen, J., & Algozzine, B. (2015). Systematic alignment of

dual teacher preparation. *Rural Special Education Quarterly, 34*(1), 30-36.

Bell, C., Gitomer, D., McCaffrey, D., Hamre, B., Pianta, R., & Qi, Y. (2012). An

argument approach to observation protocol validity. *Educational Assessment, 17*,

62-87.

Benedict, A. E., Thomas, R. A., Kimerling, J., & Leko, C. (2013). Trends in teacher

evaluation: What every special education teacher should know. *Teaching

Exceptional Children, 45*(5), 60-68.

Berliner, D.C. (2013). Problems with value-added evaluations of teachers? Let me count

the ways! *Teacher Educator, 40*(4), 235-243.

Berry, A. B. (2012). The relationship of perceived support to satisfaction and

commitment for special education teachers in rural areas. *Rural Special Education

Quarterly, 31*(1), 3-14.

Burnett, A., Cushing. E., & Bivona, L. (2012). *Use of multiple measures for

performance-based compensation.* Retrieved from

http://files.eric.ed.gov/fulltext/ED533704.pdf#page=3&zoom=auto,-21,739

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

evaluations: New Jersey's ACHIEVE NJ. *International Journal of Educational

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

Cancio, E., Albrecht, S., & Johns, B. (2013): Education and treatment of children. *A

Comparison of the Instructional Context for Students*, *36*.

Christensen, B., Jeager, M., & Loren, R. (2015) Teaching students with severe

emotional and behavioral issues. *Best Practices Guide to Intervention,*36(2),

23-56.

Coryn, C., Noakes, L., Westine, C., & Shroter, D. (2011). A systematic review of theory-

driven evaluation practice from 1990 to 2009. *American Journal of Evaluation,

32*(2), 199-226. doi: 10.1177/1098214010389321

Colorado Department of Education (2015). Practical ideas for evaluating special

education teachers. Educator Effectiveness@cde.state.co.us Retrieved October 17,

2017 from Colorado Department of Education (2015). Practical ideas for

evaluating special

education teachers. Educator Effectiveness@cde.state.co.us Retrieved October 17,

2017.

Council for Exceptional Children. (2013). *Position on Special Education Teacher

Evaluation.* Retrieved from

https://www.cec.sped.org/~/media/Files/Policy/CEC%20Professional%20Policies

%20and%20Positions/Position_on_Special_Education_Teacher_Evaluation_Back

ground.pdf

Creswell, J.W. (2014). *Research design: Qualitative, quantitative, a mixed-methods approach* (4th ed.). Thousand Oaks, CA: SAGE Publications.

Danielson, C. (2010). Evaluations that help teachers learn. *Educational Leadership, 68*(4),35–39.

Danielson, C. (2012). Observing classroom practice. *Educational Leadership, 70*(3), 32-37.

Darling-Hammond, L. (2012). *Creating a comprehensive system for evaluating and supporting effective teaching.* Stanford, CA. Stanford Center for Opportunity Policy in Education. Retrieved from https://edpolicy.stanford.edu/sites/default/files/publications/creating-comprehensive-system-evaluating-and-supporting-effective-teaching.pdf

Darling-Hammond, L. (2015). Can value added add value to teacher evaluation? *Educational Researcher, 44*(2), 132-137.

Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *Phi Delta Kappan, 93*(6), 8-15.

Dee, T., & Wyckoff, J. (2013). Incentives, selection, and teacher performance: Evidence from IMPACT. Cambridge, MA: National Bureau of Economic Research Working Paper 19529.

Derrington, M. L. (2014). Teacher evaluation initial policy implementation: Superintendent and principal perceptions. *Planning and Changing, 45*(1/2), 120-137.

Evans, B. R., Wills, F., & Moretti, M. (2015). Editor and section editor's perspective article: A look at the Danielson Framework for teacher evaluation. *Journal of the National Association for Alternative Certification, 10*(1), 21-26.

Gable, R. A., Tonelson, S. W., Sheth, M., Wilson, C., & Park, K. (2012). Education & Treatment of Children, 35(4), 499-519.

Garrett, R., & Steinberg, M. (2014). Examining teacher effectiveness using classroom observation scores: Evidence from the randomization of teachers to students. *Educational Evaluation and Policy Analysis.* doi: 10.3102/0162373714537551.

Gill, B., English, B., Furgeson, J., & McCullough, M. (2014). Alternative Student Growth Measures for Teacher Evaluation: Profiles of Early-Adopting Districts. REL 2014-016. *Regional Educational Laboratory Mid-Atlantic*.

Goe, L., Holdheide, L., & Miller, T. (2014). Practical guide to designing comprehensive teacher evaluation systems: A tool to assist in the development of teacher evaluation systems. *Center on Great Teachers and Leaders.*

Goldstein, J., & Behuniak, P. (2012). Can assessment drive instruction? Understanding the impact of one state's alternate assessment. *Research and Practice for Persons with Severe Disabilities, 37*(3), 199-209.

Hallinger, P., Heck, R. H., & Murphy, J. (2014). Teacher evaluation and school improvement: An analysis of the evidence. *Educational Assessment, Evaluation, and Accountability*, 26, 5-28.

Harris, D. N., Ingle, W. K., & Rutledge, S.A. (2014). How teacher evaluation methods matter for accountability: A comparative analysis of teacher effectiveness ratings

by principals and teacher value-added measures. *American Educational Research Journal, 51*(1)73-112

Hill, H. C., Charalambous, C. Y., & Kraft, M. A. (2012). When rater reliability is not enough: Teacher observation systems and a case for the generalizability study. *Educational Researcher, 41*, 56-64.

Holdheide, L. and Reschly, D. (2010). Challenges in evaluating special education Teachers and English Language Learner Specialists. *National Comprehensive Center for Teacher Quality*. Retrieved October 17, 2017

Holdheide, L., Browder, D., Warren, S., Buzick, H., & Jones, N. (January 2012). *Summary of using student growth to evaluate educators of students with disabilities: Issues, challenges, and next steps.* Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved from http://isbe.net/peac/pdf/using_student_growth_summary0112.pdf

Illinois State Board of Education (2000). Policy statement Least restrictive environment Illinois State Board of Education. Retrieved from http://www.isbe.net/spec-ed/pdfs/lre_policy.pdf

Individuals with Disabilities Education Act (IDEA) Amendments of 1997, P.L. 105-117, 20 U.S.C. §§ 1400 *et seq.*

Jacobs, S. (2015). Improve tenure with better measures of teacher effectiveness. *Kappan Magazine, 97*(6), 33-37.

Jones, N. D. and Brownell, M. T. (2014). Examining the use of classroom observations in the evaluation of special education teachers. *Assessment for Effective Intervention 39*(2), 112-124.

Kalogrides, D., Loeb, S., & Beteille, T. (2013). Systematic sorting: Teacher characteristics and class assignments. *Sociology of Education, 86*, 103-123.

Kane, T. J., & Staiger, D. O. (2012). Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains. Bill & Melinda Gates Foundation. Retrieved from http://eric.ed.gov/?id=ED540960

Kersting, N. B., Mei-kuang, C., & Stigler, J. W. (2013). Value-added teacher estimates as part of teacher evaluations: Exploring the effects of data and model specifications on the stability of teacher value-added scores. *Education Policy Analysis Archives, 21*(6/7), 1-39.

Kindzierski, C. M., O'Dell, R., Marable, M. A., & Raimondi, S. L. (2013). You tell us: How well are we preparing teachers for a career in classrooms serving children with emotional disabilities? *Emotional and Behavioral Difficulties, 18*(2)179-195.

King-Sears, M. E., Carran, D. T., Dammann, S. N., & Arter, P. S. (2012). Multi-site analyses of special education and general education student teachers' skill ratings for working with students with disabilities. *Teacher Education Quarterly, 39*(2), 131-149.

Kowalski, T. J., & Dolph, D. A. (2015). Principal dispositions regarding the Ohio teacher evaluation system. *AASA Journal of Scholarship and Practice, 11*(4), 3-20.

Lawson, J. (2014). Value-added modeling: Challenges for measuring special education

teacher quality. *InterActions: UCLA Journal of Education and Information

Studies, 10*(1). Retrieved from http://escholarship.org/uc/item/9r67085n

Lazarus, S. S., & Reike, R. (2013). Leading the transition from the alternate assessment

based on modified achievement standards to the general assessment. *Journal of

Special Education Leadership, 26*(1), 25-30.

Maharaj, S. (2014). Administrators' views on teacher evaluation: Examining Ontario's

teacher performance appraisal. *Canadian Journal of Educational Administration

& Policy*, (152), 1-58.

Maine Department of Education. (2012). Databases on state teacher and principal

evaluation policies. Retrieved from

www.maine.gov/doe/accountability/.../meec_7-20-12_Indicators.pdf

Marzano, R. J. (2012). The two purposes of teacher evaluation. *Educational Leadership,

70*(3), 14-19.

Marzano, R., & Toth, M. D. (2013). *Teacher evaluation that makes a difference: A new

model for teacher growth and student achievement*. Alexandria, VA: ASCD.

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach (Applied

social research methods)* (3rd ed.). Thousand Oaks, CA: SAGE Publications

McCaffrey, D. F. & Buzick, H. M. (2014). Is value-added accurate for teachers of

students with disabilities? Carnegie Foundation for the Advancement of

Teaching: Stanford, CA. Retrieved from

http://www.carnegieknowledgenetwork.org/briefs/teacher_disabilities

Minnesota Department of Education. (2014). Teacher development and evaluation for

 special education teachers. Retrieved from

 https://education.state.mn.us/mdeprod/idcplg?IdcService=GET_FILE&dDocNam

 e=054647&RevisionSelectionMethod=latestReleased&Rendition=primary

Murphy, J., & Hallinger, P. (2013). Leading via teacher evaluation the case of the

 missing clothes? *Educational Researcher, 42*(6), 349-354.

Murray, M. L. (2012). Leadership to promote inclusion: Perceptions of elementary

 principals on inclusion, so-teaching, and differentiated instruction. (Unpublished

 doctoral dissertation). University of Pittsburgh, Pittsburgh, PA.

Murray, J. (2014). Critical issues facing school leaders concerning data-informed

 decision-making. *The Professional Educator*, *38*(1), 1.

National Comprehensive Center for Teacher Quality. (2010). Special edition highlights

 resources on teacher evaluation. Washington, DC: TQ Research & Policy Briefs

National Comprehensive Center for Teacher Quality. (2010a). Special edition highlights

 resources on teacher evaluation. Washington, DC: TQ Research & Policy Update.

National Comprehensive Center for Teacher Quality. (2010b). Survey of special

 educators conducted by the National Comprehensive Center for Teacher Quality

 with support from the Council for Exceptional Children [aggregated survey

 results]. Washington, DC: Author.

National Comprehensive Center for Teacher Quality. (2011). *A practical guide to*

 designing comprehensive teacher evaluation systems.

Phillips, K., Balan, R., & Manko, T. (2014). Teacher Evaluation: Improving the process. *Transformative Dialogues: Teaching & Learning Journal*, *7*(3), 1-22.

Polikoff, M. S. (2014). *Common core state standards assessments*. Washington, DC: Center for American Progress.

Polikoff, M. S., & Porter, A. C. (2014). Instructional alignment as a measure of teaching quality. *Educational Evaluation and Policy Analysis*. doi: 10.3102/0162373714531851

Ravitch, D. (2016). *The death and life of the great American school system: How testing and choice are undermining education*. Basic Books.

Ruffini, S., Makkonen, R., Tejwani, J., & Diaz, M. (2014). Principal and teacher perceptions of implementation of multiple-measure teacher evaluation systems in Arizona (REL 2015-062). Washington, DC: US Department of Education, Institute of Education Sciences, National Center for Education Evaluation and regional Assistance, Regional Educational Laboratory West. Retrieved from http://ies.ed.gov/ncee/edlabs

Sartain, L., Stoelinga, S. R., & Brown, E. R. (2011). *Rethinking teacher evaluation in Chicago: Lessons learned from classroom observations, principal-teacher conferences, and district implementation*. Chicago, IL: Consortium on Chicago School Research. Retrieved from http://eric.ed.gov/?id=ED527619.

Sawchuk, S. (2014, April 2). Tenn. Teachers' union takes evaluation fight into the courtroom: Lawsuit calls system arbitrary, flawed. *Education Week, 33*(27), 8-9.

Sawchuk, S. (2015, September 1). Teacher evaluation: an issue overview. *Education Week, 35*(3). Retrieved from http://www.edweek.org/ew/section/multimedia/teacher-performance-evaluation-issue-overview.html#Unions

Schwartz, N. (2013). Survey captures teacher perceptions of evaluation system. Tennessee Department of Education *Classroom Chronicles*. Retrieved from http://tnclassroomchronicles.org/survey-captures-teacher-perceptions-of-evaluation-system

Schulze, R. (2014). School principal leadership and special education knowledge. *University of Massachusetts-Amherst.* Retrieved October 17, 2017 from http://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1016&context=dissertations_2

Sledge, A., & Pazey, B. L. (2013). Measuring teacher effectiveness through meaningful evaluation: Can reform models apply to general education and special education teachers? *Teacher Education and Special Education, 36*(3), 231-246.

Semmelroth C. L., Johnson E. S., & Allred K. (2013). Special educator evaluation: *Cautions, concerns and considerations.* Journal of the American Academy of Special Education Professionals.

Smylie, M. (2014). Teacher evaluation and the problem of professional development. *Mid-Western Educational Researcher, 26*(2), 97-111.

Spina, N., Buckley, P., & Puchner, L. (2014). Shifting practices in teacher performance evaluation: A qualitative examination of administrator change readiness. *Education Leadership Review, 1*(2), 113-128.

Teachscape. (2011, July 20). Charlotte Danielson's Framework for Teaching (FFT) chosen as a teaching evaluation rubric for New York state. Retrieved from https://www.prnewswire.com/news-releases/charlotte-danielsons-framework-for-teaching-fft-chosen-as-a-teaching-evaluation-rubric-for-new-york-state-125879623.html

The Bill and Melinda Gates Foundation. (2018). "What we do." Retrieved from https://www.gatesfoundation.org/What-We-Do

The Danielson Group. (2014). Special ed scenarios: Extended examples of levels of performance in special education. Retrieved from https://www.danielsongroup.org/special-education/

National Comprehensive Center for Teacher Quality. (2013a). Fixing classroom observations. Retrieved from http://tntp.org/publications/view/evaluation-and-development/fixing-classroom-observations-how-common-core-will-change-teaching

National Comprehensive Center for Teacher Quality. (2013b). Perspectives of irreplaceable teachers: What America's best teachers think about teaching. Retrieved from http://tntp.org/publications/view/perspectives-of-irreplaceable-teachers-best-teachers-think-about-teaching

U.S. Department of Education. (2010). *A blueprint for reform: The reauthorization of the Elementary and Secondary Education Act.* Washington, DC: Author. Retrieved from http:// www2.ed.gov/policy/elsec/leg/blueprint/blueprint.pdf

U.S. Department of Education. (2011). "Developing effective teacher evaluation systems: Conversation with Charlotte Danielson." *School Turnaround, 11*(1). Retrieved from https://www2.ed.gov/programs/sif/sigsummernewsletter.pdf

U.S. Department of Education, Office of Special Education and Rehabilitation Services, Office of Special Education Programs. (2011). 30th Annual Report to Congress on the Implementation of the Individuals with Disabilities Education Act (IDEA), 2008. Washington, D. C. Retrieved from http://www2.ed.gov/about/reports/annual/osep/2008/parts-b-c/30th-idea-arc.pdf

U.S. Department of Education (2016). *Teacher Compensation*. Retrieved October 17, 2017 from https://www.ed.gov/oii-news/teacher-compensation

(2016a). Federal Disability Definitions, Title 34: Education; Sec. 300.8 Child with a disability. Retrieved October 17, 2017 from https://www.ctc.ca.gov/docs/default-source/credentials/creds/federal-disability-definitions.pdf

Van Lint, J.H. (2012). *Introduction to coding theory,* 3rd ed. Eindhoven, Netherlands: Springer.

White, Cowhy, Stevens, & Sporte. (2012). *"Designing and Implementing the Next Generation of Teacher Evaluation System"*. The University of Chicago Consortium on Chicago School Research.

Whitehurst, G. J., Chingos, M. M., & Lindquist, K. M. (2014). Evaluating teachers with

    classroom observations: Lessons learned in four districts. Retrieved from

    http://www.brookings.edu/research/reports/2014/05/13-teacher-evaluation-

    whitehurst-chingos

Whitehurst, G. J., Chingos, M. M., & Lindquist, K. M. (2015). Getting classroom

    observations right. *Education Digest, 80*(7), 20-28.

Williams, J., & Dikes, C. (2015). The Implications of Demographic Variables As Related

    To Burnout Among A Sample Of Special Education Teachers. *Education*, *135*(3),

    337-345.

Widener, R.E. (2011). Evaluating Special Education Teachers: Do We Get the Job Done?

    A Regional Perspective. *East Tennessee State University*. Retrieved October 17,

    2017 from http://dc.etsu.edu/cgi/viewcontent.cgi?article=2478&context=etd

Appendix A: Position Paper and Policy Recommendation

A policy paper recommendation to the Midwest School District (MSD) administrative

board, concerning revising special education teacher evaluation systems.

Introduction

Current methodologies surrounding teacher evaluation in American primary

education are deficient in ability to evaluate educators of students with developmental or

learning disabilities, also known as special needs students. In particular, educators in the

traditional educational environment – who do not serve students with special needs – are

often subject to evaluation of their effectiveness which either relies upon standardized

testing, value-added (year-over-year) assessment, or upon a standardized rubric of

efficacy derived from the Danielson Framework for Teaching (FFT) model. Though these

methods are deficient, they are even less effective when they are applied to special

educators. In these environments, the same rubrics are applied to evaluate teacher

efficacy, but the evaluators – primarily school principals – often lack the expertise in

special education necessary to mount an effective evaluation. This effect is worse in

classrooms where special education students are taught alongside traditional students, by

two educators, where the proficiency and causal effectiveness of either educator can often

be difficult to ascertain. In essence, current special education teacher evaluation

methods/systems are not realistic to judge these teachers' performances in the classroom.

The current system of observing special education teachers in the classroom and basing

their performance off students' state test scores is not effective. This work provides a

description of this problem, supported in the literature, as they lead to recommendations

for practice, and then present a systemic approach to correcting for this core deficiency in teacher evaluation in American primary education.

The Problem

Evaluation criteria for special education teachers should not be the same as those of general education teachers. While teacher evaluation systems typically emphasize student achievement and teacher practice, but few have the demonstrable capacity to distinguish among specialty area educators, meet the challenges of accurately measuring achievement growth for students, and link that growth to teacher effect. To this end, the problem which this work evaluates is methodological in nature; At present, research indicates that methods of teacher evaluation are constrained by current methodologies, especially those by which teacher performance is predicated on the results of standardized testing, or results from specific performance rubrics, often under the FFT model, as determined by periods of classroom observation. Though the deficiencies in these models are clear, they are often based upon a shared level of understanding of classroom effectiveness between educator and evaluator, and these deficiencies can be reduced due to the experiential understanding held by principal-evaluators who were once classroom educators themselves. However, the same cannot be said for the evaluation of special educators. First, the principals conducting such evaluations – in the majority of cases – do not come from special education backgrounds, meaning that there is an essential 'disconnect' in theoretical understanding between these two actors. Due to this difference in experience and understanding, and especially with regard to the needs of special education students, the primary deficiencies of the standard assessment rubric

are shown. These must be mitigated, if these educators' performance is to be judged in an accurate manner.

The Current Policy

While teacher compensation was linked to seniority in the past, it is now commonly linked to data- and observation-driven indicators of student and teacher performance (DoE, 2016, p. 1). Such systems of evaluation, however, are only as effective as they are comprehensive, and backed by evaluator expertise regarding the classrooms (and educators) being evaluated. Research indicates that there are major deficiencies in current modes of educator evaluation in the traditional classroom, but these may be reduced (or mitigated) through shared experiences between educators and the (often former teacher) principals conducting their professional evaluation. In the case of special educators, however, principals' efficacy in evaluation is reduced – to the limited efficacy of rubric-driven evaluation, as under the FFT model – as a function of their *lack* of shared experience and knowledge with the principals overseeing their evaluation (Schulze, 2014). The implications of this disconnect has manifest in the widespread use of policies by which special educators are evaluated ineffectively. Interpretation of survey findings lead me to believe that under the current model, lack of understanding held by principal-evaluators leads to strong special educators being evaluated as poor, which may lead to a reduction in their morale, or poor special educators being evaluated strongly, which leads to special education students – who require the closest level of attention – not receiving the instruction and guidance to which they are entitled by law.

As will be shown in the research to follow, several areas of deficiency present in education evaluation policy drive the evaluation of special education teachers. In particular, the current FFT framework for teacher evaluation employs a rubric model which standardizes all teacher behavior (Garrett & Steinberg, 2015). This model has been viewed as deficient by all educators, and often as an exercise without meaning (Phillips et al., 2014). However, experiential common ground between principals and educators may mitigate many of the deficiencies of this model, as evaluators may have a strong informal understanding of what comprises a strong classroom and can apply this understanding to their rubric (Danielson, 2011).

This is not the case, however, when principals evaluate special educators, as there is often a strong difference in expertise, compounded by the differences between the traditional and special education classroom (Darling-Hammond et al., 2011; Widener, 2011; Colorado, 2015). This reduces the capacity of the principal-evaluator to properly consider (or to grade) a special educator's performance in the classroom setting. In essence, current evaluative policy regarding classroom educators fails to consider the unique circumstances faced by special educators.

In MSD in particular, special education teacher evaluation systems are not effective. Current policy at MSD is to base special education teachers' performances on test scores, like those provided to all general education students (MSD, 2015). However, many special education students, or students who receive Individualized Education Programs (IEPs) do not learn at a similar rate to general education students, and therefore cannot score as well on tests as average, general education students.

Furthermore, MSD has failed to acknowledge the special needs of students with IEPs (MSD, 2015). Special education teachers at MSD are required to participate in IEP meetings and maintain documentation of student progress. For MSD, this has recently been added in an addendum as a performance indicator for special education teacher evaluation. This addendum, however, still does not address all special education teachers' evaluation needs (MSD, 2015).

Research

A host of changes which resulted from the Education for All Handicapped Act of 1975, by which students with disabilities are required to receive educations which are "free and appropriate," with least number of "restrictions" (Widener, 2011, p. 10). The idea of *least restrictive* indicates the need to "include" children with disabilities in the traditional classroom, and only to remove them from the traditional educational environment when the use of "supplementary aids and services" cannot be achieved in a satisfactory manner (Widener, 2011, p. 14). Though such modes of education have allowed special needs students to be taught in a nonrestrictive manner, educator evaluation has suffered as a result. Principal-evaluators often lack specific training in how best to evaluate special education environments, and the educators therein, and this factor is one compounded by their "minimal guidance" with respect to how to observe special education teachers for their classroom efficacy (Widener, p. 11).

As a result, current models of teacher evaluation focus on "student achievement and teacher practice," yet fail to differentiate between 'traditional' educators and those teaching at risk populations (Holheide & Reschly, 2010, p. 2). Though teacher

effectiveness in the classroom setting is the "most influential school-based factor" which informs student achievement, teacher evaluation must recognize "diverse teacher roles and responsibilities," as when educators teach students with special needs (p. 3).

Traditional educator evaluation tends to focus upon the "personal traits, skills, and…dispositions" of the educator, to the exclusion of quality of instruction, especially as students are "enabled…to learn" (Darling-Hammond, 2012, p. i). Such evaluations have long relied upon "principal-conducted classroom observations," typically through a 'checklist' of classroom conditions and "teacher behaviors" which often fail to gauge the "quality of instruction" (Burnett et al., 2012, p. 3). This has resulted in a system which identifies only the "very worst teachers," and fails to recognize excellence, especially when such education deviates from areas where the evaluator has expertise.

Recent years have witnessed a transition to "observation frameworks" by which "comprehensive expectations" of educators are set through rubrics which outline distinct "teaching practice and professionalism" factors which educators must meet to be deemed satisfactory (Burnett et al., 2012, p. 3). Though this is an improvement over the earlier evaluation framework, it nonetheless is deficient with respect to evaluating educators of special needs and learning-disabled students (Burnett et al., p. 3).

The origin of this modern variant on educator evaluation can be traced to education reform known as 'Race to the Top' (RTTT), which included a "teacher evaluation mandate," satisfied by several "primary measures of teacher performance," including classroom observations, student achievement of "learning objectives," and "value-added scores," by which the educator's efficacy (through student performance

testing) is compared to student performance in previous years (Garrett & Steinberg, 2015,

p. 2). Under this mandate, teacher performance measures have been gauged through use

of the Danielson Framework for Teaching (FFT), through which teachers are evaluated

for their performance in planning, preparation, classroom environment, instruction, and

fulfillment of educational responsibilities (Garrett & Steinberg, p. 2). At present, the FFT

is the "only measure used in the evaluation of teachers" in grades without standardized

exams, and has also been used to evaluate special educators (p. 2). When there are no

standardized test scores are available, FFT may "account for upward of 85% of a

teacher's performance evaluation" (p. 2).

Given the extraordinary limitations in the evaluation of the traditional educator,

any method of evaluating the effectiveness of teachers of students with disabilities

requires special attention, especially due to the lack of broad evidence-based consensus

upon which special educators might be more effectively evaluated (Council, 2012, p. 2).

In particular, special educators – independent of evaluators – are often well-aware that

little is known about "whether student growth can be adequately measured for students

with disabilities and appropriately attributed to teachers for the purpose of teacher

evaluation" (Holdheide et al, 2012, p. 1).

Core deficiencies in current modes of educator evaluation, especially in

classrooms serving special needs students, pertain to the current reliance on standards-

based measures of educator effectiveness. Common use of *value-added models* in

individual teacher evaluation are typically predicated upon the idea that "measured

achievement gains" for the teacher's students provide strong indicators of a teacher's

"effectiveness" (Darling-Hammond et al., 2011, p. 1). However, this conceptualization of the teacher's role in student progress is limited by misconception, including that student learning is "influenced by the teacher alone," and can be tracked (and attributed) independently from classmate growth as well as "other aspects of the classroom context" (Darling-Hammond, p. 1). In special education, evaluative deficiency is indicated by how special educators often teach in a *tandem* manner, to students who learn along non-"at-risk" peers, meaning these students' progress can be tied to not only dynamic classroom factors, but resulting from two teachers who teach and administer classrooms in conjunction.

Moreover, for *dedicated* special educators who do not teach in a tandem classroom, additional deficiencies manifest with regard to traditional evaluation. Dedicated special educators (1) Often teach *smaller* classrooms than their 'traditional' educator peers, leading to "less reliable estimates of teachers' effects on student performance" (NCCTQ, 2011). (2) Service delivery for students with special needs typically vary greatly, thereby increasing difficulty in attributing student gains to educator performance is inhibited in accuracy and fairness (NCCTQ, 2011). Finally, if such standardized assessments given to students with special needs are "not multistage or item-level adaptive", they may fail to measure such students' growth, and by extension, the effectiveness of their educators in performance evaluation (p. 5).

Two major recommendations are presented by Darling-Hammond et al. (2012) with respect to how to improve current models of special educator evaluation. First, evaluations must include "multi-faceted evidence of teacher practice, student learning,

and professional contributions," as considered in an "integrated" fashion which incorporates the propriety of the curriculum and evidence-based expectations of student progress (Darling-Hammond et al., 2012, p. iv). Secondly, any evaluator of educator proficiency – or the lack thereof – must be knowledgeable about instruction, as well as well-trained with respect to specific student needs, especially in an integrated special education classroom, to facilitate their provision of productive feedback and support for educators' "ongoing learning" (Darling-Hammond, p. iv). Moreover, use of (3) "Multiple measures" is encouraged, by which special educators' performance measurement can be calculated in these complex classroom environments.

Widener (2011) recommended that schools add more (4) "Professional development opportunities" for school administrators, which focuses upon the differing roles and expectations of such educators (Widener, 2011, p. 57). Finally, schools must (5) "examine their evaluation and observation" methods to inform the addition or modification of instrumentation by which the "additional roles and responsibilities" special educators hold, both individually and in a tandem classroom setting, might be incorporated into evaluative criteria (Widener, p. 58).

Such criteria would consider "measures of learning" in the special education classroom, and give specific weights to such criteria, including "measures of growth sensitive enough to indicate accelerated achievement" by which achievement gaps for special needs students can be closed (Colorado, 2015, p. 20). In addition, such evaluation protocols must employ "multiple indicators" of special educator performance, including "development of [special education] lesson plans," skill in providing special education

students with "access to the general education curriculum," and the educators' ability to implement strategies for special education student instruction which are appropriate to students with disabilities (Council for Exceptional Children, 2013).

Synopsis of the Study

This review of the extant literature has considered current methodologies by which educators are evaluated for their conduct, professionalism, and efficacy in the classroom, and has linked the deficiencies in this model to the experiential disconnect between principals – who conduct such evaluations – and the special educators whom they evaluate. Previous studies reflect significant deficiencies with the current rubric-driven model of educator evaluation, but also acknowledge the shared experiential understanding between educators and evaluating principals, and present such informal social connections as a means by which such deficiencies can be overcome. However, the same cannot be said for special educators undergoing evaluation. When these professionals are evaluated under the same standard model as traditional general educators, they are often judged by evaluators who lack their experience or understanding of the special education environment. The following policy recommendations will indicate means by which this deficiency, as it manifests in lack of mutual understanding, can be overcome.

Policy Recommendation

In order to compensate for the deficiencies in special educator evaluation as a function of standardized assessment methods, there are several options which can be proposed to educational authorities. The first body of recommendations is based on two

factors which have been presented in the literature, and which amount to (a) Improving rater competency with multiple evaluators and (b) Replacing the method of student test scores with in-classroom observation.

Improving Rater Competency with Multiple Evaluators

Research findings illustrates that requiring evaluators, especially principals, to follow a set guideline or to undergo additional training increased rater reliability and improved the quality of special education teacher evaluations (Derrington, 2014; Murray, 2014). Employing multiple reviewers in various settings (including direct, explicit instruction and whole-group instruction) helps to ensure reliability and validity in special education teacher evaluation (Semmelroth & Johnson, 2014). These findings align with others that found that "multiple observations and multiple raters are critical for ensuring acceptable levels of measurement score reliability" (Semmelroth & Johnson, 2014, p. 131). Additional evidence has determined that while administrators may feel that they have the "ability to provide fair and meaningful evaluations of special education teachers," increased training is necessary (Semmelroth & Johnson, p. 6).

Because of these findings, it is recommended that MSD remove its current method of requiring untrained principals to be the sole personnel responsible for evaluating the performance of special educators. Instead, MSD should employ a multiple-evaluator method using trained evaluators.

Using In-Classroom Observation as Evaluation Method

MSD current evaluation methods for student education teachers is based solely on test scores of special education students. However, as mentioned above, many special

education students or students who use IEPs do not learn at a similar rate to general

education students, and therefore tend to not score as well on tests as average, general

education students. Instead, MSD requires that special education teachers participate in

Individual Educational Program (IEP) meetings and maintain documentation of student

progress.

Although research shows that even in-classroom observation methods may not be

the most effective method for evaluating special education teachers' performance (Jones

& Brownell, 2014), current systems, such as value-added scores are not a suitable option

for many special education teachers. To this end, observation systems appear to be a

superior option, provided that multiple raters participate in teacher evaluation

(Semmelroth & Johnson, 2014).

Accommodations for Diverse Learners

Administrators have a range of responsibilities, including ensuring that

educational accommodations and modifications are provided to diverse learners. Students

with disabilities may require changes in lesson plan or expectations, or those which assist

them to overcome or work around their disability. To this end, regardless of which

assessment tool or process an administrator uses, the evaluation of special educator is

affected. This indicates that principal-evaluators must evaluate special education teachers

in light of not just differential expectations of student success, but with respect to lesson

plans (and success) often tailored to each student.

In total, this effort might also be ensured through employing the recommendations

presented by Darling Hammond (2012), Widener (2011), Colorado (2015), and Council

(2012). Evaluators must (1) base their evaluations of teacher performance upon the differential expectations of special education student progress, as contrasted with such measures of 'traditional' general classroom students; By extension, schools must also (2) employ additional measures of performance in their evaluations, which take into account differential expectations and understandings of student performance. In order to alter their evaluation rubrics in an effective manner, it will be critical to (3) provide principal-evaluators with additional education into special education processes. If principals do not hold a similar level of expertise with respect to special education students – and the responsibilities and expectations of special educators – there is a strong likelihood that they will fail to evaluate these educators in an effective manner.

To fulfill these recommendations, it will be necessary to approach school administrators and present them with core deficiencies in the methods they employ to evaluate special education teachers, especially as such deficiencies are supported by the extant research. Improving evaluator proficiency in the special education classroom will form a primary goal. Principals who evaluate such classroom educators must receive additional education to mitigate the current experiential gaps which preclude them from evaluative accuracy. If this goal is untenable, then a proposal will be presented by which such evaluative rubrics will be altered to account for requirements of the special education classroom.

Recommended Course of Action

Though the direct funding of special education-specific evaluation rubrics may result in superior *efficiency* in special educator evaluations, this will require a significant

investment. To this end, superior evaluator education (as through in-service or other classes) must be presented as a *simple* approach to ensuring that evaluators can apply the same informal awareness of classroom requirements which they bring to their evaluation of the traditional classroom, as through a pure evaluator education designating positive excellence. Such an infusion of expertise will be framed as a path to these educators becoming better acquainted with the classroom factors of their special educators, without having to alter a broad swath of evaluative standards and rubrics overnight. Evaluators' expertise (and role in reducing the drawbacks of standard evaluation) must be enunciated, and the proposal will center upon broadening their skill-set, so that they can bring the same informal expertise to bear in their evaluation of their special educators.

The school district may not be receptive to these change recommendation proposals, no matter how sensible they seem. As bureaucracies tend to be resistant to the *imposition* of change, this outsider proposal may be met with significant internal resistance. Thus, this proposal must be presented as a *simple* and *sensible* option which will require some (but not a major level of) investment, either of time and money. Direct lobbying for these change measures will thus focus upon the fact that they will not necessitate direct changes to written standards and rubrics.

However, because this proposal involves the direct solicitation of internally-funded change efforts in a public bureaucracy, public pressures may be brought to bear if no receptivity to change – or its necessity – is received. Through contacting news-media, this issue may be framed in terms of *child welfare*, mainly of vulnerable special education students, and in terms of their right to receive an education despite physical or

developmental disabilities. Thus, bureaucratic recalcitrance in the face of minimal

proposed additional evaluator education, so framed, may result in sufficient media

pressure 'embarrassing' the bureaucracy by bringing attention to this issue, and the

bureaucrats resisting change to educator evaluation policy.

Project Evaluation

Because the special education classroom is a complicated environment which

differs widely from the traditional environment in many ways, it will be necessary to

employ *educator* expertise in the course of evaluating the effectiveness of this change

proposal. Provided that the evaluators and administrative bureaucracy are receptive to the

proposed change and choose to fund direct continuing education of their principal-

evaluators to enhance their informal expertise regarding special educator requirements,

educators so evaluated will form the crux of understanding with respect to whether this

policy proposal has been effective in its goals. As all special educators in this school

district will continue to be evaluated along the same rubric and test-driven standards as

employed in the traditional classroom, these educators' perceptions of their evaluators

expertise in conducting such evaluations and mitigating their deficiencies, as determined

by internal survey, will provide a key data-driven indicator of success. If special

educators believe that their evaluators have *gained* such expertise sufficient to mitigate

common failings in traditional evaluation, then this policy change will have succeeded in

its goal.

Conclusion

A wide body of extant research shows that methodologies of evaluating educators are deficient, particularly in light of the mandates for such evaluation since the passage of 'No Child Left Behind' in 2002, and 'Race for the Top' in 2010. Both laws carried strong mandates for educator performance evaluation. However, evidence indicates that shared traditional classroom experience between evaluating educators and teachers can mitigate some core identified deficiencies in educator evaluation; In particular, because most evaluating principals were once traditional classroom educators, they have a clear sense of what constitutes an effective classroom and can employ this knowledge toward even a constrained rubric-driven evaluation.

As has been considered, evaluating professionals (mostly principals) often lack a specific special education background; This factor necessarily limits their effectiveness in evaluating special educators, whose roles and responsibilities, as well as the 'metrics of performance among their students, differ considerably from those which are presented in the traditional classroom. This position paper has presented evidence to indicate that principal-evaluators must not be impeded in such evaluations by their lack of experience. Research evidence has been presented, toward recommendations by which such 'gaps' in experience – which stand to damage special educator morale or student service delivery can be mitigated. In essence, this work has shown that evaluation rubrics must be altered to reflect the differential expectations of the special education classroom, and principal-evaluators must receive further education to bolster their understanding of these critical differences.

The preceding recommendations for policy – and evaluation of such policy – center on a range of core concepts in elementary education, particularly the likelihood of these institutions to become *stagnant*, or likely to adhere to a normative "status quo," over time, even if the methodologies which inform that status quo are deficient. As a result, though this work has outlined core means by which educators can be better-evaluated, through increasing *evaluator* expertise regarding the special education classroom. As has been shown, informal systems of awareness and expertise are often sufficient to result in a comprehensive evaluation which mitigates many of the core deficiencies of common evaluative mechanisms in the traditional classroom, but evaluators often lack the awareness of expertise upon which to mitigate such deficiencies in special education. Through increasing the level of evaluator education regarding special education, superior evaluations of such educators can be attained.

References

Burnett, A., Cushing. E., & Bivona, L. (2012). *Use of multiple measures for*

performance-based compensation. Retrieved from

http://files.eric.ed.gov/fulltext/ED533704.pdf#page=3&zoom=auto,-21,739

Colorado Department of Education (2015). Practical ideas for evaluating special

Education Teachers. Retrieved from

https://www.cde.state.co.us/practicalideassped

Council for Exceptional Children. (2013). *Position on special education Teacher*

Evaluation. Retrieved from://www.cec.sped.org.

Danielson, C. (2011). Evaluations that help teachers learn. *Educational Leadership 68*(4):

35-39.

Darling-Hammond, L. (2012). Creating a comprehensive system for evaluating and

supporting effective teaching. Stanford, CA. Stanford Center for Opportunity

Policy in Education.

Darling-Hammond, L., Beardsley, A., Haertel, E., & Rothstein, J. (2011). Evaluating

Teacher Evaluation: What we know about Value-Added Models and Other

Methods. *Phi Delta Kappan.* Retrieved October 17, 2017 from

https://eml.berkeley.edu/~jrothst/publications/Evaluating%20Teacher%20Evaluat

ion%20-%20PDK%20prepublication.pdf

Derrington, M. L. (2014). Teacher evaluation initial policy implementation:

Superintendent and principal perceptions. *Planning and Changing, 45*(1/2), 120-

137.

U.S. Department of Education (2016). *Teacher Compensation*. Retrieved October 17, 2017

Garrett, R., & Steinberg, M. (2014). Examining teacher effectiveness using classroom observation scores: Evidence from the randomization of teachers to students. *Educational Evaluation and Policy Analysis*. doi: 10.3102/0162373714537551.

Holdheide, L. and Reschly, D. (2010). Challenges in evaluating special education Teachers and English Language Learner Specialists. *National Comprehensive Center for Teacher Quality*. Retrieved October 17, 2017 from http://www.gtlcenter.org/sites/default/files/docs/July2010Brief.pdfHoldheide et al, 2012

Jones, N. D. and Brownell, M. T. (2014). Examining the use of classroom observations in the evaluation of special education teachers. *Assessment for Effective Intervention 39*(2), 112-124.

Murray, M. L. (2012). Leadership to promote inclusion: Perceptions of elementary principals on inclusion, so-teaching, and differentiated instruction. (Unpublished doctoral dissertation). University of Pittsburgh, Pittsburgh, PA.

National Comprehensive Center for Teacher Quality (2011). *A Practical Guide to Designing Comprehensive Teacher Evaluation Systems.* Retrieved October 17, 2017 from http://www.lauragoe.com/lauragoe/practicalguideevalsystems.pdfSemmelroth & Johnson, 2014

Schulze, R. (2014). School principal leadership and special education knowledge.

University of Massachusetts-Amherst. Retrieved October 17, 2017 from

http://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1016&context=dissert

ations_2

Widener, R.E. (2011). Evaluating Special Education Teachers: Do We Get the Job Done?

A Regional Perspective. *East Tennessee State University.* Retrieved October 17,

2017 from http://dc.etsu.edu/cgi/viewcontent.cgi?article=2478&context=etd

Appendix B

Survey Instrument

Do you agree or disagree that the following affect principals' perceptions of effectiveness of the evaluation of special education teachers as compared to general education teachers? How strongly?	Strongly disagree	Disagree	Agree	Strongly agree

Please feel free to add any details on any of your responses below.

- Q2. Teacher evaluations must differ based on the differential roles and responsibilities of special educators.
- Q3. Educator assessment lacks precision when the evaluator does not have knowledge of specific practices which contribute to improvements in student performance.
- Q4. Teaching special needs students requires more than just good instructional practice.
- Q5. Educating students with special needs demands strategies which include social-emotional skills development, through which students may be aided in mediating between social connections, both in and outside of school.
- Q6. Grade level state

assessments may be misaligned with special needs students' abilities.

- Q7. Special needs students must be expected to achieve the same education goals as 'general content' students at his/her grade level.
- Q8. Teaching functional skills are viewed as wasted time if they don't improve test scores and can negatively influence teach evaluation.
- Q9. The same student performance and student evaluation results should be used to assess teaching quality for both special and general educators.
- Q10. Teacher performance analysis is most meaningful and helpful when constructed on several rankings and well-defined teaching standards.
- Q11. Validated evaluation measures based on widely accepted standards are essential.
- Q12. A single test score cannot accurately represent teacher effectiveness.
- Q13. Teacher assessments must employ a range of measures of teacher efficacy, including portfolios and students' weekly assessments
- Q14. Classroom

> observations are the only useful source of information for measuring educator performance.
> - Q15. The distinguished ranking in teacher evaluation is largely associated with grade level student growth.

Do you agree or disagree that the following key indicators of teaching quality apply to teacher assessment? How strongly?	Strongly disagree	Disagree	Agree	Strongly agree

> - Q16. One may identify the influence of a special education students' other educators using teacher ratings.
> - Q17. Special educators are evaluated as distinguished with the same frequency as general education teachers.
> - Q18. Teacher evaluation is objective when the evaluator has experience in teaching and assessing students with special needs.
> - Q19. Feedback is considered authentic, meaningful, and relevant to special education teachers when the evaluator has prior teaching experience with special needs students.

Do you agree or disagree that the following present barriers to conducting teacher evaluations?	Strongly disagree	Disagree	Agree	Strongly agree

How strongly?

- Q20. I do not have the prerequisites (qualifications, experience etc.) to conduct special education teacher evaluations.
- Q21. There is a lack of employer support.
- Q22. There is inadequate time to prepare.
- Q23. The assessment tools are inadequate.
- Q24. Inadequate school budget and resources.
- Q25. Government regulation and policy
- Q26. High workload or workplace responsibility

Principal Demographic Information and School Profile
- Q27 Are you male or female?
  Female
  Male

- Q28. How old are you as of your last birthday?
  26-30
  31-35
  36-40
  41-45
  46-50
  51-55
  56-60
  61-65
  66 or older

- Q29. Have you have completed formal education? To what level?
  Bachelor's degree
  Master's degree
  Doctoral degree or equivalent (Ph.D., Ed.D., J.D., M.D.)

- Q30. Do you have a teaching degree?
    Yes
    No

- Q31. How many years of experience have you had as a general educator?
    0
    1-5
    6-10
    11-20
    More than 20

- Q32. What is the current enrollment in your school?
    1-499 students
    500-999 students
    1,000-1,500 students
    1,501-1,999 students
    2,000 students or more

- Q33. What percentage of students in your school have special needs?
    0
    1-5
    6-10
    11-19
    20 percent or more

- Q34. What percentage of teachers in your school are special educators?
    0
    1-5
    6-10
    11-19
    20 percent or more

- Q35. What regional location do you represent?
    Central
    Far South Side
    North/Northwest Side
    South Side
    Southwest Side
    West Side

- Q36. What goals would you like your special educator evaluation process to address?

_____
_____
_____
_____
_____

- Q37. What are the strengths of your special educator evaluation process?

  _____
  _____
  _____
  _____
  _____

- Q38. What are the weaknesses of your special educator evaluation process?

  _____
  _____
  _____
  _____
  _____

- Q39. What changes, if any, would you like to make to your special educator evaluation process?

  _____
  _____
  _____
  _____
  _____

- Q40. Please use the following space to add any comments you may have about special educator evaluation practices and related issues.

  _____
  _____
  _____
  _____
  _____