School of Education Student Capstone Projects   School of Education

Summer 2017

Educator Perceptions Of A Teacher Evaluation Systems

Erin N. House
Hamline University

EDUCATOR PERCEPTIONS OF A TEACHER EVALUATION SYSTEM

by

Erin N. House

A capstone submitted in partial fulfillment of requirements for the degree of Master of Arts in Teaching.

Hamline University

Saint Paul, Minnesota

August 2017

Primary Advisor: Susan Manikowski
Secondary Advisor: Patrick Tanis
Peer Reviewer: Meg Nord

TABLE OF CONTENTS

CHAPTER ONE

Introduction

Most teachers I know are extremely passionate about their profession, myself included. We become teachers because we love kids and we love learning, and we want to inspire kids to love learning, too.

But how does passion translate into practice? Quantifying a teacher's efficacy is a challenging undertaking, one that is clouded with variables and fraught with politics. Today, many teacher evaluation systems in American schools are largely intended to be tools that enable teachers to reflect on and improve their practice in order to best instruct their students. In the Midwest state where I teach, the Educator Effectiveness model has been used since 2014 to evaluate teacher performance, with the stated goal "to improve the education of all students in the state…by supporting guided, individualized, self-determined professional growth and development of educators" (Educator effectiveness online training, 2016).

That is the theory behind the evaluation model, but I am curious to know how educators feel about it, whether they believe it is a meaningful way for them to enhance their teaching skills. My research question is the following: What are teacher perceptions of the Educator Effectiveness system of evaluation? Since the intended outcome of the evaluation system is for teachers to use it as a growth and development tool, the usefulness of the system is predicated on whether teachers believe it is a necessary and effective model for them to improve their professional practice.

In this chapter I will explain my early experience with evaluation systems as a Teach for America corps member. Next, I will briefly outline the current model in my state, how it has been implemented in my current school district, and my own personal experiences with it. Finally, I will set the context for my research by sharing my beliefs and perceptions of teacher evaluation, including its potential benefits, as well as the role it could play in recruiting, retaining, and developing highly effective educators.

As a teacher who is highly invested in the performance of my students, colleagues, and school, I believe it is of critical importance to determine whether this state-mandated teacher evaluation program is having its proposed effect within our walls. I can say from experience that the Educator Effectiveness model requires a great deal of energy, planning, and paperwork, and so I feel it is important to determine whether the outcomes of the program are worth the work put into it.

My Early Experience With Teacher Evaluation

An Analogy for Evaluation

One of my favorite educators of all time is Fred Rogers – yes, that Fred Rogers, of *Mister Roger's Neighborhood*. As a young girl, the gentle, loving messages he shared in his children's television show were a daily reminder to be brave and curious, and to always be myself; lessons that I have applied throughout my life. While Mister Rogers, of course, never had a show on educator evaluation, I can apply some wisdom to the topic from his 1977 book, *Having An Operation*. On the final page of the story, Mister Rogers states:

> There are some things that are hard to do – like having an operation. But there are
>
> people who care about you who can make these things seem easier. Doing

something that's hard can make you grow – and you can feel proud of the way

you are growing. (Rogers, 1977, p. 16)

As a new teacher, being observed and evaluated by administrators often felt akin to

having an operation – a scary, unavoidable procedure that was meant to make you better.

But it was hard. Teaching is unique in that it is simultaneously a very public, yet very

personal endeavor, and I certainly did not enjoy being scrutinized while trying to conduct

this difficult job. Yet, like Mister Rogers said, the process seemed easier when it was

clear that my evaluator cared about me as a developing educator. To open myself up to

critique is still difficult, but I believe it is ultimately the best way to grow and improve.

Teach for America

Perhaps this attitude comes from the fact that I have never known any differently.

I came into teaching through Teach for America (TFA), a non-traditional teacher

preparation program. Teacher evaluation is an integral component of TFA's teacher

training model. The very first time I ever stood before a classroom of students at my

TFA Summer Institute in Los Angeles in 2009, I had an array of peers and coaches

observing, recording, and critiquing my performance.

During the six weeks I spent teaching summer school in an L.A. charter school,

there never was a time that I did not sit down with a veteran teacher after a lesson to

break down what went well and what I could improve on, using the TFA Teaching As

Leadership rubric. This was the normal, accepted routine of being a corps member, and

one that I knew would continue during my two-year commitment to the program.

After the summer experience concluded, I taught third grade at a charter school

for low-income boys in North Minneapolis as a brand new TFA corps member. I taught

full time under the close supervision of Teach for America staff, while also working on

my teacher certification courses at Hamline University.

Additionally, my school had its own system of teacher evaluation, one that was

very high stakes – one tied to retention, salary, and merit pay bonuses. It was an intense,

challenging time. I was a new teacher. I was adjusting to a new city. I was

overwhelmed with work and coursework. My school placement was difficult and

demanding. I was constantly under a microscope, both by TFA and my school

administrators.

A Mentor's Encouragement

There were many times when I wanted to quit. I remember a meeting where I sat

down with my TFA Program Director, who served as both my coach and evaluator, and

told her that I didn't think I could handle the pressure and didn't have what it takes to

make it through the two-year commitment. I thought that my students deserved better.

My coach almost laughed as she reassured me that in all the years she had spent

mentoring new teachers she had a good idea of which corps members could hack it, and

which couldn't, and that she knew for a fact that I belonged to the former group. She also

said I was the best possible teacher for my students, because my commitment to them

would drive me to develop my practice in order to provide them the best education I

could.

That vote of confidence from someone who had spent so much time in my

classroom watching me struggle in my new role was so reassuring, and also quite pivotal

for me. I not only stuck out my two TFA years, but ending up stayed teaching for two

more years in Minneapolis, and ultimately chose teaching as my vocation and permanent profession now in my hometown.

Mister Rogers was right. Difficult things – like teacher evaluations – can ultimately make us stronger. I am proud of the way I have grown and continue to grow since I first walked into that classroom in L.A. almost eight years ago.

The Educator Effectiveness System

In 2013 I moved home to teach at a school I attended as a child, and now am under a new evaluation system. A 2011 State Senate bill mandated that the Department of Public Instruction (DPI) "develop an educator effectiveness evaluation system" for public school principals and teachers (2011 Senate Bill 461). According to the legislation, half of the evaluation score must come from student performance measures, and half must come from core teaching standards (2011 Senate Bill 461).

In 2014, therefore, DPI adopted a new teacher evaluation system called Educator Effectiveness (EE). It is based on Charlotte Danielson's model of evaluation, which uses four domains to gauge a teacher's efficacy: planning and preparation, the classroom environment, instruction, and professional responsibilities (Danielson, 2007, p. 1). These domains are broken into several components with detailed descriptions.

The Evaluation Cycle

In the district where I teach, teachers are put on a three-year cycle, where the third year is a "summary year" (The educator effectiveness user guide, 2014, p. 9). During the summary year, teachers engage in in a self-review of their practice at the beginning, middle, and end of the year, and set individual Student Learning Objectives (SLOs), which they track and monitor throughout the year.

Teachers are also formally and informally observed by their administrator three to five times that year, and are evaluated on each of the several components in Danielson's four domains. Based on their evaluation results and progress on their SLOs, teachers receive a final evaluation score at the end of the year on a 1-4 performance scale, where Level 1 is unsatisfactory, Level 2 is basic, Level 3 is proficient, and Level 4 is distinguished (Educator effectiveness user guide, 2014, p. 8).

High-Stakes Decision-Making

One key question of any teacher evaluation system is how the teacher performance ratings will be used throughout the district and state. In the pilot phase of Educator Effectiveness, DPI barred participating districts from using the scores for "high-stakes decisions, like awarding merit pay or termination" (Luders, 2013). This might not always be the case, however. According to Luders (2013), DPI sees Educator Effectiveness mainly as a tool to support and develop struggling teachers, yet explicitly leaves the door open for the system to possibly someday "inform the full range of human resource decisions." While teacher tenure, salary, and retention are not currently tied to performance ratings in the state, there remains the potential.

District Implementation

As stated above, Educator Effectiveness became the law in public schools in the state in 2014. When it was first implemented in my school district in 2014, the Superintendent expressed hope for the new model in an interview with the local paper. She said she believed that EE would promote professional growth for educators, allowing teachers to choose for themselves what they want to work on by writing their own Student Learning Objectives (Wachter, 2014). She also thought it would encourage

collaboration, enabling teachers to work together to meet their goals, and to have conversations with principals about instruction (Wachter, 2014).

In 2015, a year after EE was implemented in my school district, the district staff was asked to complete a satisfaction survey. According to Dohms (2015), overall satisfaction in the district had dropped by 14% in the two years since the last survey in 2013. Two common themes from the survey were complaints of not having enough time to complete professional responsibilities, and a decrease in the quality of working conditions. The district Director of Teaching and Learning cited Educator Effectiveness online tools among other factors that "negatively influenced the quality of [teachers'] working conditions" (Dohms, 2015).

Educator Effectiveness in Practice

When I came to the district in 2013, I was considered a "probationary" teacher (despite having four years of experience teaching in Minneapolis), and was therefore evaluated under the system that was in place at the time. When Educator Effectiveness was enacted in the state in 2014, I was automatically put on a summary year for that school year, and again in 2015 until my three-year probationary status was over. Therefore, I have a great deal of personal experience under the new EE model, having been through the evaluation process two years in a row. In fact, this past school year wass the only year in my eight-year career that I have not been observed and evaluated by my administrator.

I found the EE process relatively painless. At first, since it was a new system, there was a lot of learning as far as what the process involved: paperwork, due dates, and how to use the online platform. Our school district dedicated much of our professional

development time to explaining the procedures and requirements, however, which was very useful. Also, I am blessed with a very understanding, patient, and completely non-threatening principal who is always a joy to have visit my classroom. She is very encouraging and highly complimentary of my teaching. Finally, I naturally enjoy and engage in self-reflection, so it was easy for me to set goals, work towards them, and track them along the way under the EE model. I would have done that anyway, and in fact, still follow this process, despite not being on a summary year.

The biggest frustration with the system for me was the lack of organization. The DPI website that teachers were supposed to use to work through and submit much of our paperwork was rolled out too soon and was riddled with problems, and ultimately had to be shut down, forcing us all to complete hard copies of our documents. My principal was constantly giving us updates on changes to deadlines and procedures, which caused a lot of confusion. I recognize that it was a new system and bound to have some glitches, but it certainly was stressful to try to be responsible and stay on top of things under such conditions.

My other critique of the system is its lack of "teeth." When I finished my final evaluation meeting with my principal at the end of the first year, I remember thinking, "That's it?" I received a nice score, had a nice talk with my nice principal, and went on with my day. It was all rather anti-climactic. Maybe that is a good thing, and maybe other educators had different experiences, but after a year of working towards this big goal, and being observed and scored on my practice, I expected more feedback on areas of improvement.

Perhaps if this score was tied to compensation, or if teacher leadership positions were determined by it, it would carry more weight. I'm not necessarily saying I would like any of that to happen at this point, but I'm merely reflecting on the fact that the process seemed to me like another hoop that the state bureaucracy set out for me to jump through.

My Perceptions of Evaluation

My Personal Evaluation Philosophy

I have learned to see evaluations not as an occasion to be perfect, but rather to be honest. I try to make my observed lessons as accurate a reflection of my practice as possible, and because I am a devoted and skilled professional, I have no reason to fear an extra set of eyes watching my students and I conducting our normal business.

Of course, I would be lying if I said that my heart did not beat a little faster for the first five minutes after an evaluator walks through the door, but once I settle down and remember to rely on my planning and delivery, I relax into the routines and rapport I have developed with my students.

When it comes time to meet with my evaluator for my post-observation meetings, nothing she has to say is ever that big of a surprise to me. Because I am constantly reflecting on my own practice as a natural part of my professional approach to my job, I know my areas of strength and weakness. What can be powerful in those feedback conversations, however, is receiving advice about what I could have done to make my lesson more effective.

Potential Benefits of Evaluation

I believe that evaluations that don't result in a helpful, growth-oriented conversation about specific aspects of teaching afterwards serve little to no purpose. Some of my evaluators over the years have been more useful than others in providing meaningful feedback. I have found that the people best able to give constructive criticism are those who have themselves taught in a similar educational setting for several years.

Unfortunately, I have had multiple evaluators who have no real understanding of the context of my teaching, because they have never taught the same grade level or subect themselves, and therefore are not able to offer many helpful ways to improve my skills. I believe that a key qualification for any administrator or coach to observe and evaluate a teacher should be that that person has also served a similar position for 5-10 years, for only then are they able to understand and critique my planning, methods, and delivery.

Another aspect of evaluation that I believe is potentially very powerful is peer observation. I had the opportunity to participate in this at a KIPP charter school in Minneapolis that I briefly taught at before relocating to my hometown. Teachers there were encouraged and required to go into each others' classrooms to not only provide feedback, but to get ideas on ways to improve their own teaching. While a peer's eyes can be more unnerving than an administrator's, I think that when done right, this practice helps produce a sense of collegiality, trust, and continuous learning among educators.

As teachers, we sometimes exist as islands. While the current Professional Learning Community (PLC) trend in education helps to combat this in areas of planning and data analysis, I think that learning from each other's classroom management, procedures, and instructional delivery can be immensely helpful. This is true for new and veteran teachers. At my current school, some open-minded, highly experienced teachers

have borrowed some of my own practices in their own classrooms, so that they can stay innovative.

In general, I believe that the more open educators are to feedback and continuous learning, the better. Granted, not all approaches and techniques fit everyone's individual teaching style, but I believe that a learning community that is founded on a commitment to reflection, growth, and mutual trust for *all* the stakeholders involved – students, teachers, and administrators alike – will ultimately lead to better outcomes. In my experience, feedback that comes from a reliable source and is focused on teacher development and student achievement is the most productive and almost always appreciated.

The Need for Research

After reflecting on my experiences with evaluation, the Educator Effectiveness model in particular, I'm curious to know what my colleagues' opinions are. In theory, the evaluation system is meant to improve teacher practice and inspire teachers towards independent growth and development as educators. I wonder if this is happening, or if EE is merely putting another burden on teachers, causing them more stress and diverting their time away from activities that would enhance their practice. If we all agree that the ultimate goal of teacher development is better learning outcomes for students, we need to ensure that our current systems actually support that.

In my opinion, one of the best ways to drive student achievement is to recruit and retain top candidates into the teaching field. I have often heard that with the new expectations and requirements being put on teachers in the state, education is becoming even less of an attractive profession than ever before, and fewer college students are

pursuing it as a career. I wonder if new state initiatives like EE will ultimately have the reverse effect and reduce the number of new, quality teachers joining the field.

I believe there needs to be a way to continue to encourage educators to reflect on how their teaching impacts student performance that does not prove to be an insult to their professionalism or a burden on their time. In recent years, teachers have been demonized and blamed for a whole host of problems, especially in my state. I don't believe teaching as an island is a productive model, but teachers need to feel supported and respected for being the professionals they are, who do a difficult job under increasingly difficult circumstances.

I hope my research gives educators an opportunity to share their perceptions of Educator Effectiveness. I want to learn what, if anything, they find helpful from the system, and learn about their ideas for improvement. I would love to have a larger conversation about what they believe the purpose of educator evaluation is and how it fits into their practice. Does it enhance or diminish our profession? What do they believe is the future of teacher evaluation in our state? Hearing directly from the individuals it impacts will be a good test of the system's utility in the state.

Conclusion

My state is currently in the third year of a new teacher evaluation system that has already transformed the educational landscape in the state, and has the potential for leading to even more drastic changes for educators in the future. This is a timely and important topic, and one that requires a thoughtful, honest look at the implications of such reforms on all stakeholders involved.

In this chapter I have discussed my own experience with teacher evaluation, including my early years as a Teach for America corps member, and my current years under the Educator Effectiveness system. I have summarized what Educator Effectiveness entails for teachers, and shared my own opinions and perceptions of the system. I concluded the chapter by sharing the reasons for my research and the goals of my project.

In the next chapter, I will review the current literature on educator evaluation. I will begin by giving an overview of the accountability movement in education, which has led to evaluation reforms, including Educator Effectiveness. I will then detail different goals and approaches to teacher evaluation, considering the benefits and criticisms of each. Finally, I will present findings from research of teacher perceptions of evaluation systems across the nation, in order to set the stage for my own research on teacher perceptions of Educator Effectiveness.

CHAPTER TWO

Literature Review

My research question is: What are teacher perceptions of the Educator

Effectiveness system of evaluation? Because the evaluation system is so new to my state,

there is not much current literature on how it has been received by the teachers it impacts.

I seek to determine whether or not teachers believe the evaluation process in the state is

achieving the intended purpose of improving teacher development and increasing student

achievement.

This chapter reviews the relevant literature supporting my research question. In

the first section, I will provide a brief history of the accountability movement in

education in order to explain the conditions that set the stage for the Educator

Effectiveness (EE) model of evaluation. The second section considers different goals of

and approaches to teacher evaluation, because the EE model employs some of these to

evaluate teachers. The third and final section looks at current research into teacher

perceptions of evaluation systems across the country, as well as implications for potential

future reforms. Teacher perception is the focus of my research into the EE model, so

reviewing the results of current research into other models is essential for the

development of my own study.

History of the Accountability Movement in Education

Introduction

The past three decades in American education have been a whirlwind of reform. As society becomes increasingly complex, the needs of the nation's students also become more diversified, and policymakers, scholars, and school districts are attempting to find solutions to meet those needs.

This section traces the various reform movements throughout United States history. The first part describes the important publication *A Nation at Risk*, and the second part moves to early attempts at standards-based reforms. No Child Left Behind is the focus of the third section, followed by more accountability reforms, such as recommendations from *The Widget Effect* and value-added measures. The final part describes the Race to the Top grant competition and its implications for education reform.

This context is an important backdrop for my research question, as it helps explain the different aspects of the Educator Effectiveness system, and why the state decided to implement it.

A Nation at Risk

One of the earliest and most influential federal documents calling for public school reform was the 1983 report *A Nation at Risk* by the National Commission on Excellence in Education. Then president Ronald Regan's education secretary, Terrel H. Bell, authorized the report in an effort to publicize the perceived dismal state of public education in America. Warning of a "rising tide of mediocrity," the authors lamented the poor performance of American students compared to other nations on international tests (National Commission on Excellence in Education, 1983). They claimed that if allowed to continue, this trend would result in the loss of manufacturing jobs to other nations such

as Japan, South Korea, and Germany, and thus force the U.S. to lose its competitive

economic edge (National Commission on Excellence in Education, 1983).

Indicators of mediocrity. In addition to the poor performance of American youth

on international assessments, the authors of *A Nation at Risk* identified students' rising

rates of illiteracy, lower SAT scores, and lower scores on other standardized tests as

further indicators of the failure of the U.S. education system. The authors pointed out

that these reduced achievement rates occurred at the same time that technology was

becoming more complicated, prevalent, and requiring greater skill and intelligence from

America's youth to create, build, and operate it (National Commission on Excellence in

Education, 1983).

Causes. *A Nation at Risk* also criticized public schools for having low

expectations of students, and described an educational landscape in which several

students chose easy electives over rigorous academic courses, and spent less time on

school work than other industrialized countries. The authors blamed curricula in schools,

as well, which they claimed had become "homogenized, diluted, and diffused" over the

years (National Commission on Excellence in Education, 1983). Another problem the

Commission addressed was the fact that American teachers came from the lowest quartile

of college classes, and spent too much time on methods courses during their training.

Teacher salaries were low in comparison to other nations, and there was a severe shortage

of math and science teachers (National Commission on Excellence in Education, 1983).

Recommendations. To address these problems, the report included several

recommendations. It called for a lengthened school day, as well as for improvements to

teacher preparation programs and higher teacher salaries. In terms of accountability, it

recommended stronger graduation requirements for students, as well as "more rigorous

and measurable standards, and higher expectations, for academic performance and

student conduct" (National Commission on Excellence in Education, 1983).

  Impact. *A Nation at Risk* was a lively piece of rhetoric, at times employing

hyperbolic language to persuade readers of the mediocrity of American students. In a

segment regarding the perceived threat of the U.S. being overtaken in "education

attainments" by other nations, the report states, "If an unfriendly foreign power had

attempted to impose on America the mediocre educational performance that exists today,

we might well have viewed it as an act of war." (National Commission on Excellence in

Education, 1983). Such imagery left its mark on education stakeholders. According to

the Center on Education Policy (2008), *A Nation at Risk* was a "seminal event" for public

education in the nation, initiating conversation and debate in policy circles about raising

standards for student achievement and teacher performance (p. 17). The need for

accountability in public schools, that is, to systematically monitor student performance,

became a key focus for several policymakers following the publication of *A Nation at*

Risk (Center on Education Policy, 2008).

Early Standards-Based Reform Attempts: America 2000, NCEST, and Goals 2000

  After *A Nation at Risk* warned of the degradation of educational quality in

American schools, reformers turned their attention to the creation of more rigorous

academic standards that go beyond basic facts and skills. In line with the central tenant

of *A Nation at Risk*, this was an attempt to keep America competitive with other nations.

According to the Center on Education Policy (2008), reformers also were focused on

extending better learning opportunities to *all* students, especially disadvantaged ones, to

ensure equity in education (p. 18). Another component of these early reform efforts was to change achievement measures, creating assessments that require complex thinking. The Center on Education Policy (2008) describes the reformers' desire to create "tests worth teaching to," including performance-based assessments such as portfolios, open-ended questions, and hands-on projects (p. 19).

America 2000. The first time the federal government dabbled with involvement in standards-based reforms came in 1989 during President George H.W. Bush's Education Summit with state governors in Charlottesville, Virginia. Again, the main purpose of the summit was to ensure the U.S. would stay competitive with other countries by increasing student achievement. It also hoped to create greater educational uniformity among the states. The participants came up with six goals, which later translated into Bush's education strategy called America 2000 (Department of Education, 1991).

The Center on Education Policy (2008) reports that America 2000 aimed to develop world-class standards and national tests for students, but made this voluntary for states (p. 19). The strategy never became law itself, but parts of it were present in 1994's Goals 2000: Educate America Act, President Clinton's signature education legislation, which will be covered later in this section.

National Council on Education Standards and Testing. Another important milestone in standards-based reform during President Bush's administration was the establishment of a group called the National Education Goals Panel, which was tasked with overseeing progress made towards the education goals developed by the federal government and the states (Center on Education Policy, 2008). This body created the National Council on Education Standards and Testing (NCEST) to work on developing

national standards and assessments. In a 1992 publication entitled *Raising Standards for American Education*, the authors write, "In the course of its research and discussions, the Council concluded that high national standards tied to assessments are desirable" (National Council on Education and Testing [NCEST], 1992, p. 2). The authors claimed that reading and math skills in American public schools were low-level, and that low expectations pervade education due to the lack of high expectations for students and teachers (NCEST, 1992, p. 2).

The Council believed that having high national academic standards would lead to high expectations and serve several functions in American society: promote equity in education, make the U.S. more competitive economically, give an increasingly diverse and mobile population a shared set of values and knowledge, and "preserve democracy and enhance the civic culture" (NCEST, 1992, p. 3). The Council called for the creation of national, not federal, standards that reflect high, not minimal, student competency. These standards were not intended to provide a national curriculum, but rather give direction. The Council also stated that the national standards would not be mandated, but rather voluntary for states to implement, and would be dynamic and open to changes or revisions (NCEST, 1992, p. 3).

Debate stymies reform. According to the Center on Education Policy (2008), the voluntary nature of these national standards was the downfall of their development and implementation. Stakeholders argued over how prescriptive and specific they should be. Some called for them to be broad and used merely as a guide for teachers to develop their own curriculum, while others wanted them to be specific and provide for no local

curriculum discretion. There were also debates between behaviorists and constructivists on the approach to teaching and learning, and where the national standards fit in (p. 23).

Goals 2000. When Bill Clinton came into power in 1993, Congress began drafting new education legislation, based in part on the goals of Bush's America 2000. Under Clinton, Goals 2000: Educate America Act became law in 1994, its defined purpose being to "provide a framework for meeting the National Education Goals" by, in part, "promoting coherent, nationwide, systemic education reform" (Goals 2000: Educate America Act, 1994). This legislation also established the National Education Standards and Improvement Council that would monitor the yet voluntary national content and student performance standards. This move towards standards-based reform can be seen as the precursor to what was to come in No Child Left Behind legislation in the very next presidential administration.

No Child Left Behind

In 2001, George W. Bush became president and began pushing forward his agenda on education policy in the early days of his administration. The No Child Left Behind Act was passed by Congress in 2001 and signed into law by President Bush in January 2002 (No Child Left Behind Act [NCLB], 2002). This new law constituted a federal commitment to the standards-based reform movement that had been developing in the country since *A Nation at Risk* was published in 1983.

Landmark school reform legislation. The ambitious, long-term goal of NCLB was 100% proficiency in math and reading for every student in the country by 2014. Notable items in the new act included mandatory yearly standardized testing of all 3rd through 8th grade students in math and reading, as well as a new standardized measure of

Annual Yearly Progress (AYP) set by individual states to see if students were on track to reach proficiency. Local Education Agencies (LEAs) were tasked with identifying schools that were not making adequate progress and take action to correct them. States and LEAs were required to report school performance and teacher quality to parents and the public on an annual school report card (NCLB, 2002).

"Failing schools." NCLB required the restructuring of schools not making AYP after one year of correction. Some restructuring options available to LEAs included reopening the failing school as a charter, replacing all or most of the staff (including the principal), creating a contract of operation with a proven public or private company, or having the state take over school operations (NCLB, 2002).

Varying levels of proficiency. A key component of NCLB was that states developed their own assessments, as well as student proficiency levels. According to Goldstein (2014), this led to a great deal of variation between states regarding the rigor of both the tests, and what was considered passing (p. 185). She cites Texas as an example, which set 13% as its proficiency level on its state standardized tests. Massachusetts, however, set high, rigorous standards for its students, which resulted in fewer students being deemed "proficient" (Goldstein, 2014, p. 186).

Criticism of NCLB. Critics claim that NCLB not only failed to accomplish the ambitious goal of 100% proficiency by 2014, but it also resulted in several other negative effects on education. Ravitch (2013) claims that states spent hundreds and millions of dollars and up to 20% of instructional time preparing for and taking the required annual tests (p. 13). She sees this as a waste of educational time and resources, and one that directly benefits private testing companies who develop such training and assessment

tools (Ravitch, 2013, p. 12). Goldstein (2014) describes how many schools were forced

to narrow their curricula to focus more time on math and reading, the tested subjects in

NCLB, and abandon other subjects such as social studies, science, art, music, physical

education, and recess (p.187). She also explains how teachers spent more time on test-

prep with the so-called "bubble kids," medium-ability students with a better chance of

passing standardized tests than their lower-performing peers, which resulted in some of

the neediest children being left behind. Finally, Goldstein reports that there were several

documented examples of cheating and misreporting in this new era of high-stakes

standardized testing. In order to reach proficiency targets, some schools would either

suspend struggling students the day before the test, or simply tell them to stay home on

testing day (Goldstein, 2014, p. 187).

A new education landscape. No Child Left Behind represented a realization of

many aspects of the standards-based reform movement. Schools were now held

accountable for student achievement, yet the high-stakes nature of the law had some

unintended consequences. The Center on Education Policy (2008) says that after NCLB,

the importance of standardized test scores led many teachers to now use test results to

inform their teaching, and not the actual standards for learning, resulting in what the

authors term "test-based reform" (p. 29).

More Reforms: *Widget Effect* and Value-Added

As discussed above, the No Child Left Behind legislation of 2001 was a major

victory for standards-based reformers, but even more reforms were to come. While

NCLB focused primarily on *school* improvement, the focus now shifted to individual

teachers.

The Widget Effect. In 2009, the New Teacher Project published a report called

The Widget Effect: Our National Failure to Acknowledge and Act on Differences in

Teacher Effectiveness. The introduction to the publication, lays out the report's general

thesis: "A teacher's effectiveness – the most important factor for schools in improving

student achievement – is not measured, recorded, or used to inform decision-making in

any meaningful way" (New Teacher Project, 2009, p. 1). The report claimed that public

schools in the U.S. do not differentiate between levels of teacher quality. School

evaluation records for teachers indicate that each teacher in school districts across the

country is doing a great job. The authors say there is little to no data on which teachers

are the most and least effective, or in their words, schools "fail to distinguish great

teaching from good, good from fair, and fair from poor" (New Teacher Project, 2009, p.

2).

The New Teacher Project calls this the "Widget Effect": districts and

administrators making the assumption that the effectiveness level of each teacher is

roughly the same. The report looked at four states – Arizona, Colorado, Illinois, and

Ohio – and determined that teacher performance is only taken into account for

remediation and dismissal of teachers, but not for recruitment, hiring and placement,

professional development, compensation, tenure, retention, or layoffs (New Teacher

Project, 2009, p. 2). The problem with this, according to the report, is that "excellence

goes unrecognized" and "poor performance goes unaddressed" (New Teacher Project,

2009, p. 4).

The report blames outdated and ineffectual evaluation systems and under-trained

administrators for not recognizing teacher differences, and claims that this pervasive

phenomenon is highly destructive to U.S. public education: "In its denial of individual strengths and weaknesses, it is deeply disrespectful to teachers; in its indifference to instructional effectiveness, it gambles with the lives of students" (New Teacher Project, 2009, p. 2).

Recommendations. The report calls for changes to be made in the practice and use of teacher evaluations. It recommends that districts "adopt a comprehensive performance evaluation system that fairly, accurately, and credibly differentiates teachers based on their effectiveness in promoting student achievement" (New Teacher Project, 2009, p. 5). Administrators need to be trained how to use the new evaluation system effectively, and be held accountable for doing so. The final two recommendations are the most controversial – tying performance evaluations to high-stakes policies, such as teacher assignment, compensation, retention, and dismissal, and removing barriers from dismissing teachers if they fail to improve. The report claims that in order for the evaluations to be meaningful and rigorous, they have to be tied to real consequences (New Teacher Project, 2009, p. 6).

Value-added measures. The value-added model of teacher effectiveness is another reform aimed directly at teacher performance, intended to accomplish exactly what the authors of *The Widget Effect* call for. According to Ravitch (2013), this statistical model was developed by William Sanders in Tennessee to differentiate effective from ineffective teachers, with the goal of getting rid of so-called bad teachers (p. 100). Goldstein (2014) explains that the value-added model looks at the progress students make on standardized tests every year, and determines whether or not they exceed expectations on each end-of-the-year standardized test, based on the predictions

from the previous year's score (p. 205). This allegedly shows the value of the teacher

over the course of the year, where student learning is treated as a finite quantity, and the

teacher is treated as the variable (Ravitch, 2013, p. 100).

Closing the gap. Education reformers claim that students who have three to five

consecutive "effective" teachers (as determined by value-added measures) will show

enough improvement over the course of those years to effectively close the achievement

gap. Conversely, they believe that students who have ineffective teachers will continue

to fall farther and farther behind (Ravitch, 2013, p. 101). According to Ravitch (2013)

reformers like Stanford economist Eric Hanushek believe that the cure for improving

public education is to rank teachers, from high to low, based on the test scores and gains

of their students, and fire the bottom 5-10%. They want teacher evaluations to be

overhauled using test-based, value-added measures to identify and reward effective

teachers, and "deselect" the lowest performing teachers (pp.104-105).

This topic will be revisited in the following section of this chapter, including

critiques of the implementation and use of value-added measures in determining teacher

effectiveness.

Race to the Top

Value-added measures were a key component of President Obama's Race to the

Top competitive grant. Race to the Top, developed by then Education Secretary Arne

Duncan, was a part of the American Recovery and Reinvestment Act (ARRA), the

economic stimulus package signed into law on February 17th, 2009 (U.S. Department of

Education [Dept. of Ed.], 2009). In addition to the $95 billion earmarked in ARRA to

keep teachers employed and schools running, $4.35 billion was set aside to fund the Race

to the Top grant competition between states, rewarding states for pushing forward a reform agenda and improving student outcomes (p. 1).

Goals. The Executive Summary of Race to the Top describes the aim of the program as encouraging schools to implement "innovative strategies that are most likely to lead to improved results for students, long-term gains in school and school system capacity, and increased productivity and effectiveness" (Dept. of Ed, 2009, p. 1). The four core reform areas were standards and assessment; data systems to measure growth and inform school decision-making; teacher recruitment, development, rewards, and retention; and turning around low-performing schools (Dept. of Ed, 2009, p. 1).

Student growth in evaluations. The Executive Summary explains that states' applications were judged by a point system based on six areas, one being "Great Teachers and Leaders," which looked for states that were "improving teacher and principal effectiveness based on performance" (Dept. of Ed, 2009, p. 9). This is where *The Widget Effect* and value-added methods come into play. Race to the Top required states to use student growth on standardized tests as a portion of teacher evaluations, and for districts and administrators to use these evaluations to inform high-stakes decisions such as teacher professional development, compensation, promotion, retention, granting tenure, and dismissal if teachers failed to improve their effectiveness (Dept. of Ed, 2009, p. 9).

Teacher accountability under Race to the Top. This demonstrates the shift in federal education policy from school-level to teacher-level accountability. Ravitch (2013) writes, "Many teachers were disheartened by No Child Left Behind, which overemphasized standardized testing. Obama's Race to the Top proved even more discouraging than NCLB because it directly targets teachers as the source of student

success or failure" (p. 116). Indeed, Race to the Top addressed the same perceived

problems with teacher effectiveness laid out in *The Widget Effect*. Gottlieb (2014) writes

that Duncan felt that administrators don't do a good job of rating teachers as effective or

ineffective, so states needed to "fix" their evaluation methods to include student

achievement data from standardized tests (p. 23). According to Gottlieb (2014), Duncan

believed this student test data should have real consequences. He blamed states and

unions for blocking this, because in his opinion, it would pave the way for great teachers

to truly make an impact on student outcomes (p. 23).

Goldstein (2014) explains how Duncan pushed this reform agenda through not

only to the states who won grants through Race to the Top, but across the country, by

using the economic recession to his advantage. She says the grant program had an

"ingenious design" by holding out "an irresistible carrot – federal funding – and directed

financially starving states to compete against one another to grasp it" (Goldstein, 2014, p.

214). According to Goldstein (2014), only nineteen states were awarded grants under

Race to the Top, but two-thirds of all states changed state laws regarding their public

school teachers so they could enter the competition. Additionally, one half of the states

who applied decided to use test scores in their evaluation of teachers (p. 214).

Impact of Race to the Top. Like NCLB, Race to the Top had several

consequences that changed the educational landscape in American public schools.

Ravitch (2013) discusses how this federal program, much like NCLB, opened the door

for private businesses to get involved in public education, as they now advise districts on

necessary services in the new educational reality, including how to redesign teacher

evaluation systems, train administrators to evaluate teachers, and optimize data-driven

instruction, (p. 15). Goldstein (2014) describes that under many new evaluation systems, many principals now have to evaluate each teacher each year, resulting in a change in their professional role and a much higher workload with immense amounts of paperwork (p. 225). She also reports more systematic cheating in districts, including the highly publicized scandal exposed in 2012 in Washington, D.C. under reformer Michelle Rhee's chancellorship, and another infamous one in Atlanta in March 2013, which included 35 teachers and administrators (Goldstein, 2014, pp. 226-227). For better or worse, NCLB and Race to the Top have resulted in high-stakes conditions for districts, principals, and teachers.

Conclusion

American public education has changed a great deal since the publication of *A Nation at Risk* in 1983. Education reformers have by and large prevailed, and influenced public policy and federal legislation to the extent that many states and districts use accountability measures for their public schools and teachers.

The United States is currently in the early days of a new administration. What public education will look like with Trump as president and his Education Secretary, Betsy DeVos at the helm of the Department of Education remains to be seen. Will they choose to continue the reform agenda established by the Bush administration and advanced by the Obama administration? Will the corporate reform model continue to drive education policy, or is there something new on the horizon for American public schools? Millions of stakeholders wait in anticipation for the answers to these questions.

This first section of the chapter examined the history of the accountability movement in order to set the stage for my own research into one specific reform: teacher

evaluation. The next section of this chapter looks specifically at new approaches to teacher evaluation that have been developed during the education reform era.

<div align="center">**Teacher Evaluation Systems**</div>

Introduction

For many years, teachers viewed the process of evaluation with a mixture of annoyance and apathy. Oftentimes, experienced teachers felt principal observations were another bureaucratic piece of business that fulfilled a meaningless requirement, a ritual that had long outlived its usefulness. Teachers and principals would play the game for the half hour observation, then the paperwork would be completed and submitted, and the whole procedure would be forgotten until it was time for the next requisite evaluation.

While initially intended to assess educational quality, the classroom observation had increasingly become an inauthentic "dog-and-pony show" performance that was in large part not a true reflection of teacher practice. Burton, Carper, and Wilburn (2011) look at teacher evaluation systems through an anthropological lens, identifying the "culture" that surrounds the traditions between the teacher and evaluator, where teachers receive a standard, expected rating, and the principal provides habitual, useless comments (p. 24). Neither party seemingly are engaged in the process or the outcome of the practice.

The first part of this section looks at criticisms of the traditional system of teacher evaluation. The next part explains the apparent paradox of teacher evaluation and ways to reconcile it. The final two parts will dive into specific reforms suggested to improve teacher evaluation, such as the value-added method and multiple measures.

This section directly relates to my research question, since the Educator Effectiveness model replaced a traditional model that was considered to be ineffective at determining teacher quality and leading to professional growth for educators. The current EE model employs several of the suggested reforms included in this section.

Critique of Traditional Systems

Traditional systems of educator evaluation have come under increasing scrutiny in recent years. As described in the previous section, school reformers have been pushing for greater individual teacher accountability in public education, which they claim will lead to better instructional outcomes and life opportunities for students. Although *The Widget Effect* of 2009 brought a great deal of national attention to the problems in teacher evaluation, education experts were calling for reform even before then.

Ineffective evaluation systems. Danielson and McGreal (2000) describe a dismal state of affairs regarding traditional teacher evaluation systems, saying that they rely on old, outdated criteria that do not get at the heart of what really matters in teaching (p. 3). Like *The Widget Effect* authors, the authors claim that too many teachers are considered "outstanding" because teacher rating scales are imprecise, and that "good teaching" is not clearly defined between the teacher and the evaluator (Danielson & McGreal, 2000, pp. 4-5). They also take issue with the hierarchical nature of traditional evaluations, in which the observer watches the teacher teach, gives a score, and gives feedback that may or may not be helpful. They claim this is because many school administrators are not expert evaluators themselves and may lack knowledge about the specific grade levels or content areas they are evaluating, which the authors claim

"undermines the evaluation process, contributing to the perception that it has little value" (Danielson and McGreal, 2000, p. 6).

Toch (2008) agrees that there are not many credible evaluation systems that truly measure teacher quality, due to lack of accountability and "staffing practices that strip school systems of incentives to take teacher evaluations seriously," by which he means tenure and union protections (p. 32). He criticizes the tendency for school districts to view credentials as indicators of effectiveness, and laments that evaluations systems don't weight instructional quality or student learning more heavily (Toch, 2008, p. 32). Toch (2008) refers to traditional evaluations as "drive by" glimpses into the classroom, consisting of a quick principal visit that includes checking discrete behaviors off of a checklist and quickly labeling the teaching satisfactory or unsatisfactory, which he feels is not doing anything to actually improve teaching and learning (p. 32).

Student achievement as measure of quality. Later critics (post-*Widget Effect)* take their analysis a step further. In addition to claiming that traditional evaluation systems don't give enough specific information to help teachers improve and that many principals just give most of their teaching staff satisfactory ratings, Marzano and Toth (2013) decry evaluation systems that don't tie student achievement to evaluation ratings (p. 3). They applaud the efforts of Race to the Top-style evaluation systems that mandate the use of student growth to determine teacher impact, along with rigorous measures of teaching skills (Marzano & Toth, 2013, p. 4).

Jackson and Remer (2014) also explain that the traditional measures of teacher credentials, such as a bachelor's degree, a state license, and proof of basic competency in the subject matter area are no longer sufficient to guarantee teacher quality (p. 1). The

authors agree that the shift in educator evaluation needs to be on student achievement

measures, claiming that "high-quality teacher evaluation data can also be used to inform

policies across the education system," including teacher preparation programs,

performance-based compensation, professional development, and equal access to

effective teachers for all students. (Jackson & Remer, 2014, p. 1).

Formative Versus Summative Evaluations

A major conflict in educator evaluation is determining the purpose of such

ratings. Danielson and McGreal (2000) describe two purposes that are often at odds with

each other: formative and summative (p. 8). Formative evaluations are intended to

provide feedback leading to the development of professional educators. The goal is to

unite all staff around student achievement, and encourage excellent teaching (Danielson

& McGreal, 2000, p. 8). Educators tend to favor this model, for it takes into account the

complex nature of teaching and is focused on mastering a very challenging profession.

Summative evaluations, on the other hand, are intended to screen out poor teachers,

leading to "legally defensible evidence" of bad teaching and potential dismissal

(Danielson & McGreal, 2000, p. 8). Legislators and policymakers favor this approach,

for it provides a clear measure of accountability to their constituents, who provide tax

dollars to public education.

Reconciling the two purposes. Danielson (2001) cites the same apparent

incompatibility between professional development and quality assurance, describing the

conflict between coaching and evaluation (p. 13). She seeks to merge the two through a

series of recommendations for improving evaluation systems, including differentiating

evaluation for novice and experienced teachers, requiring teachers to take an active role

in their own development, and conducting more "professional conversations" in the context of evaluation, leading to reflection and mutual learning among colleagues (Danielson, 2001, p. 14). As far as student achievement data goes, Danielson (2001) says that if states choose to use it to evaluate teachers, they must ensure the equity and reliability of the information, because many factors influence student learning (p. 15).

Danielson and McGreal (2000) also discuss the need to focus on student outcomes, although they make a point to say that this does not necessarily just mean standardized test data (p. 19). The authors believe that student performance should inform the picture of teacher evaluation, where educators "work backwards" from student achievement data and hold coaching conversations about the learning that is going on in the classroom. All work needs to be linked back to measurable student learning goals, in order to enhance instruction, not judge it (Danielson & McGreal, 2000, p. 19).

Evaluation Reform: Student Achievement and Value-Added Methods

Others are more adamant about including student achievement data in the actual evaluation scores of teachers, and this seems to be the direction school reform efforts are taking. Jackson and Remer (2014) explain that since the Race to the Top incentives of 2009, an increasing number of school districts are evaluating teachers more often, developing performance classifications with multiple levels, and using multiple measures of teacher effectiveness, including student achievement scores (p. 2). The authors write that only 15 states required annual evaluations and student achievement measures included in those evaluations in 2009, but by 2013, 28 states had yearly evaluations, and a whopping 41 states took student achievement scores into account when determining teacher ratings (Jackson & Remer, 2014, p. 2)

Value-added measures. Marzano and Toth (2013) say that reformers want to use student achievement scores, specifically value-added growth scores, to identify poor performing teachers, and use this information to grant or deny tenure (p. 5). Those in favor of this practice claim that students who have teachers who produce high value-added scores have far better life outcomes, including college attendance, increased salaries, and living in better neighborhoods as adults (Marzano & Toth, 2013, p. 5).

As explained in the preceding section of this chapter, value-added models (VAM) of student achievement analyze students' performance on a standardized test, using an expected growth trajectory, in order to show the contributions a teacher made to student learning (Jackson & Remer, 2014, p. 3). Marzano and Toth (2013) summarize VAM as showing "how much a student has learned since some designated point in time," with the purpose of measuring the effectiveness of a teacher (p. 5)

Criticisms of value-added measures. Despite the initial excitement over VAM, they have been found lacking in several respects. The problem with them, according to Goldstein (2014), is that the error rate for value-added measures is up to 35% when only one year of test score data is used to determine teacher effectiveness. Marzano and Toth (2013) agree, saying that VAM is inconsistent, showing huge changes for teachers from year to year, and that the results differ when different tests or statistical measures are used to calculate them (pp. 6-7) Additionally, as Jackson and Remer (2014) report, VAM cannot measure the effectiveness of every teacher in a district, since teachers who team-teach cannot have their value measured in this way, nor can teachers who do not teach in a testing grade (p. 4).

Ravitch (2013) also critiques VAM, because teachers are not the sole variables in student learning. Social scientists claim family background, especially income level, has a much greater role on a child's performance in school than teachers do. Economists estimate that differences in tests scores can be attributed to 60% family influence, and only to 20-25% school factors (pp. 102-103). Darling-Hamilton, Amrein-Beardsley, Haertel, and Rothstein (2012) echo this criticism, saying many factors affect student achievement that a teacher has absolutely no control over (p. 8). These factors include school-level ones, such as class size, curriculum, time in the school day and school year, and access to specialists and other resources; home supports and challenges; student-level, such as ability, health, attendance; peer culture and achievement; past school experience; and summer slide (Darling-Hamilton et al., 2012, p. 8)

Finally, the impact of the value-added method currently remains hypothetical. According to Goldstein (2014), the claim that the achievement gap will be closed if a student has three to five effective teachers in a row has not been tested, and even if it was, the results would not conclusively prove this was all the result of teacher influence, because value-added gains fade over time and are unstable from year to year (p. 207). Ravitch (2013) expresses a similar sentiment in her book, saying that reformers' claims that consecutive "great" teachers have the power to close the achievement gap for students "remains a theory based on speculation, not evidence," since it's never been proven to work in any school district, even where the local education culture is "fully supportive of the corporate reform faith and without a teachers' union to stand in the way" (p. 106).

Evaluation Reform: Multiple Measures

Due to the many issues with VAM described above, several reformers are now calling for the use of multiple measures in educator evaluations. These multiple measures typically include both student achievement data of some kind, as well as a more holistic look at teacher practice. Jackson & Remer (2014) claims that an effective evaluation system that informs both teacher development and accountability needs to use multiple measures to differentiate the effectiveness of teachers (p. 3). They say that a combination of student achievement measures, teacher observation, and student survey is the best, because it shows different aspects of teaching and learning, is more fair, and better informs professional development (Jackson & Remer, 2014, p. 3).

Measuring the daily classroom experience. Marzano and Toth (2013) call for the next generation of evaluations to improve the accuracy of teacher evaluations, and they provide six recommendations to achieve this, including measuring both teacher practice and student growth in multiple ways that is reflective of the daily teaching and learning experience (p. 13). They also think it is important to use evaluation to improve teacher skills by providing specific supports to struggling educators. Additionally, they feel principals and district leaders should be evaluated to determine how much support they provide teachers for professional growth and development (Marzano and Toth, 2013, p. 14).

Feedback for improvement. Darling-Hamilton et al. (2012) agree that giving teachers timely, helpful feedback from observations is instrumental in improving evaluation systems (p. 13). The authors promote a method of evaluation and development called Peer Assistance and Review, in which expert mentor teachers coach novice and experienced teachers who are struggling, providing them with evaluation and

support, but also following due process procedures for potential dismissal if teachers fail

to improve (Darling-Hamilton et al., 2012, p. 14).

Comprehensive evaluation. Finally, Toch (2008) highlights another model

called the Teacher Advancement Program (TAP) that takes multiple measures into

account when evaluating teachers (p. 32). TAP uses extensive evaluation for

instructional improvement, using the results for "coaching, career ladders, and

performance-based compensation" (Toch, 2008, p. 33). In addition to classroom

observation, the evaluation process includes multiple measures of teacher performance

such as portfolios, videos, reflections, evidence of collaboration and parent involvement,

and content-check essays, which provide a more comprehensive insight into teacher

practice (Toch, 2008, p. 33-34). Similar to the Peer Assistance and Review program

described above, TAP employs multiple evaluators, including teams of observers,

mentors, as well as peer review (Toch, 2008, p. 33-34). Although such methods can be

costly, Toch believes that they are worth it because investing in teacher development

makes educators feel like they are valued professionals who do a meaningful, important

job, and will serve to draw other young talent into the field of teaching (2008, p. 37).

Conclusion

Teacher evaluation used to be a meaningless protocol that often neither principals

nor teachers took seriously. Nearly every teacher would score "satisfactory," and that

score would be filed away for the year without much thought, until it was time for the

next compulsory evaluation.

Once this process was exposed for what it was – an ineffective way of tracking

the quality of educators and the development of their skills – reform efforts to change and

improve teacher evaluation took off. Despite the initial temptation to rank teachers based on standardized test scores, most reformers today are calling for a more measured approach that tries to account for the complex nature of teaching in new evaluation systems that are designed to help teachers reflect and improve upon their practice.

It remains to be seen how impactful these new systems will be in improving teacher quality throughout the country. The next section of this chapter will look at how teachers perceive the effect these new efforts have on the development their professional practice.

<div align="center">**Educator Perceptions of Evaluation**</div>

Introduction

In recent years, many traditional systems of teacher evaluation across the country have been overhauled to provide more teacher accountability. Reforms such as more frequent evaluations, the use of student achievement scores in teacher ratings, and tiered performance rating criteria are sweeping across the nation in the wake of Race to the Top. Supporters of such changes contend that they will lead professional growth and increased effectiveness for educators, and improved performance for students.

Whether or not teachers have bought into these new systems and believe that they fulfill their intended purpose is a key point in determining the value of new evaluation measures. Teachers, who are most directly affected by evaluation reform, have much to say about the impact of these new policies on their practice and their students' learning.

This section will examine current research into educator perceptions of teacher evaluations. The first part will present findings from recent studies that look into both

teacher and principal perspectives on evaluation systems. The final section will consider

implications for future policy decisions and further research.

My research question also asks about teacher perceptions of evaluation,

specifically the system in my state. Reviewing research that has been done in other states

into evaluation systems will prove a useful means of comparison for my own study.

Teacher Perceptions of Evaluations Around the Country

Several doctoral dissertations in recent years have focused on the topic of teacher

perceptions of new educator evaluation systems across the country. In states such as

Tennessee, Ohio, and New Jersey, researchers are looking into how teachers who are

rated under reformed evaluation systems believe the new methods impact their practice.

Value-added measures in Tennessee. Previous sections of this chapter have

explored value-added measures of student achievement and their use in evaluating

teachers. Darling-Hamilton et al. (2012) report the results of a survey of Tennessee

teachers who voluntarily were evaluated by and had their salaries tied to value-added

measures (p. 12). After three years of this system, 85% of the teachers felt that these

measures did not take important aspects of teaching into account, and 2/3 of the teachers

surveyed felt that the methods used to evaluate them were not able to distinguish between

effective and ineffective teachers (Darling-Hamilton et al., 2012, p. 12).

Job satisfaction in Ohio. Ohio is an example of a state experiencing teacher

evaluation reform. Downing (2016) surveyed 290 K-12 Ohio teachers evaluated through

the new Ohio Teacher Evaluation System (OTES), a model that includes both student

growth and observation in determining teacher ratings (pp. 51-53). The researcher

wanted to know if there was a correlation between the evaluations and teacher job

satisfaction, now that the OTES model includes student achievement data (Downing, 2016, p. 95). She found that there was no relationship between the two variables from her quantitative survey (Downing, 2016, p. 97). However, the open-ended comment section of her instrument revealed some negative perceptions of evaluations, with several teachers saying that the OTES growth measures are not equitable, because not all grades or subjects have standardized test data to use as a measurement, so these teachers get to develop their own assessments (Downing, 2016, pp. 99-100). Downing (2016) theorizes that the reason these negative perceptions did not impact the survey data about job satisfaction is because the discontent with the new evaluation system is not strong enough yet to impact general job satisfaction (p. 101).

TEACHNJ. Another state now using student achievement data to rate teachers is New Jersey. Callahan and Sadeghi (2015) explain that when New Jersey was awarded $38 million in Race to the Top funding in 2011, they chose to use the money to pilot and develop TEACHNJ, a new educator evaluation system that included four levels of teacher ratings, linked student data to those ratings, and made tenure harder for teachers to earn (p. 47). As of 2013, all educators in New Jersey were evaluated every year, with 20% of their rating coming from student growth, and 80% from an evaluation of "teacher practice" based on classroom observations (Callahan & Sadeghi, 2015, pp. 47-48).

Change in perception. Callahan and Sadeghi (2015) gave two surveys, one pre-TEACHNJ in 2012 to 254 teachers, and one post-TEACHNJ in 2014 to 364 teachers, to gauge teacher perceptions of evaluation (p. 50). They found that from 2012 to 2014, more teachers agreed or strongly agreed that more teachers got dismissed for poor performance under the new evaluation system (Callahan & Sadeghi, 2015, p. 53). The

participants also said that by 2014, teachers are observed more often, but the helpfulness of the observations decreased, because the evaluations became formulaic and procedural, with principals more focused on entering data in a computer than actually observing the lesson (Callahan & Sadeghi, 2015, p. 56). Callahan and Sadeghi (2015) concluded that TEACHNJ "turned what was once an organic, albeit infrequent process into a scripted one," and that teachers under the new system are "demoralized, and one of the contributing factors is the emphasis on rating teachers" (p. 57).

Impact on teaching and learning. Wacha (2016) also studied teachers in New Jersey to find out to what extent TEACHNJ would improve teaching and learning in public schools there (p. 6). Her methods included not only a survey, but also an open-ended questionnaire and follow-up interviews with ten teachers in a high school in New Jersey (Wacha, 2016, p. 19). Wacha (2016) reports that the teachers she studied felt the evaluation process under TEACHNJ did not improve the teaching and learning at their school, because the feedback they received from observations was neither helpful nor specific (p. 27). They also said that the professional development and support they received from their principals did not help them improve as teachers, because it was "based on educational trends" and not on their own needs (Wacha, 2016, p. 28).

Charlotte Danielson model in New Jersey. Moss (2015) conducted a qualitative study that looked at teacher perceptions of evaluations in a New Jersey high school that uses the Charlotte Danielson Framework (p. 81). He interviewed fifteen teachers from eight content areas about their experiences with evaluation (Moss, 2015, p. 84). According to Moss (2015), the school had used the Danielson model since 1997, one year

after it was published, and was a highly successful school that had won several achievement awards (p. 87).

Moss (2015) asked teachers how the Danielson Framework helped them reflect on their teaching and "improved the quality of professional conversations" with their principal (p. 93). He also asked teachers to report about the value of the feedback they received from their principal at their post-observation conferences (Moss, 2015, p. 93). According to Moss (2015), the teachers agreed that their evaluations and the feedback they received from them both helped guide their professional development and helped their administrator make decisions about staffing (p. 157).

However, teachers also cited the purpose of evaluations to be about "compliance and accountability" twice as often as they did about professional growth (Moss, 2015, p. 157). Moss (2015) identifies this as a disconnect in teachers' perceptions, with less than half of the teachers in the school seeing professional development as the purpose of their evaluations, despite their contention that the feedback from their principal was helpful (p. 157). According to Moss (2015), teachers at the school still require messaging from administrators that the focus of evaluations is on teacher development (p.158).

Principal Perceptions of Evaluations

Principals often find themselves in the difficult position of being both a judge and a coach for their teaching staff, and must walk a fine line between demanding improvement and offering support. As evaluators of teachers, principals are in a unique position to offer feedback about the value of evaluations in improving teaching and learning in their schools, and their perceptions can inform how effective evaluation systems are at providing both accountability and development for their teachers.

Comparing teacher and principal perceptions. Sheppard (2013) looked at both teacher and principal perceptions of the evaluation system in three rural school districts in southeast Georgia, surveying 227 teachers and 12 principals (pp. 5-6). According to Sheppard (2013), most teachers rated the evaluation system above average or very high in quality. This was higher than the principals' ratings, which rated the system as being only average (p. 60). In terms of feedback about evaluations, Sheppard (2013) found that the teachers believed it was adequate, and that the ideas and suggestions they received were above average in quality. Principals, however, thought the feedback was only average in its specificity and quality (p. 61). More principals than teachers felt that a large amount of time was spent on the evaluation process, as well (Sheppard, 2013, p. 61). Both the teacher and the principal groups said the purpose of evaluations was more for teacher growth than accountability, with no principal citing accountability as the purpose (Sheppard, 2013, p. 62).

Performance rating criteria. One common evaluation reform that many states are now employing is the use of tiered performance ratings for teachers. Bullis (2014) studied principals' perceptions of how teacher performance ratings affect teacher growth and effectiveness among their staff (p. 1). He looked at principals' experiences with such ratings in Florida and Massachusetts in order to inform principals in Illinois, who had recently adopted a new evaluation system in 2011. As in the other two states, the performance ratings for teachers in Illinois that went along with the new evaluation system had four levels: excellent, proficient, needs improvement, and unsatisfactory (Bullis, 2014, pp. 1-2). Bullis (2014) wanted to know what the intended and unintended

impacts of those performance ratings are on teachers, and what lessons principals in Illinois could learn from the experiences of those in Florida and Massachusetts (pp. 2-3).

Bullis (2014) based his research on survey results of the 4,533 principals in Florida and 1,854 principals in Massachusetts (p. 84). He found that while the perceived intent of the performance ratings is to promote teacher growth, the unintended result of them is low teacher morale and interference of growth (Bullis, 2014, p. 160). Principals said that teachers tend to focus more on their final evaluation score than the growth they are making, which causes them stress and can lead to burn-out (Bullis, 2014, p.162). Thus, according to Bullis (2014), another unintended impact of the performance ratings the principals cited is that teachers who are resistant to change quit or retire early. Some principals believed this is healthy for the profession, while others said that even good teachers can feel discouraged by the evaluation process and will leave teaching to pursue other work (p. 163).

Recommendations for Future Evaluation Reform

Based on the results of the studies presented above, educators and administrators have their doubts about whether new evaluation reforms have achieved their intent of improving both teacher effectiveness and student learning. Several researchers and authors have opinions on ways to further enhance and improve teacher evaluation, valuing educators as professionals while taking into account the difficult nature of teaching.

The problem with test-based accountability. The push to base all or part of a teacher's rating on student achievement scores from standardized tests has been a key component of modern education reform. Nuñez (2015) describes the modern focus of

evaluation reform based on "test scores and conformity to externally imposed standards" as a prime example of what she calls "teacher bashing and teacher deskilling" (p. 174). She says that teachers are deemed incompetent when their students fail to "achieve uniform measures of accountability," and they are no longer trusted to understand and respond to their students' needs and interests (Nuñez, 2015, p. 174).

According to Goldstein (2014), other voices in education fear that evaluation based on high-stakes test scores could create too much competition in a field that requires and benefits from collaboration (p. 210). She cites Randi Weingarten, former president of the United Federation of Teachers in New York City, as saying that the best use of value-added growth measures would be to use them for *collective* reward for a school that sees its test scores rise. If the scores go up, then all teachers at the school would get the same bonus, no matter what subject or grade level they taught.

Goldstein (2014) also talks about the benefits of evaluators focusing more on feedback and coaching than on relying solely on student achievement data to rate and rank teachers, quoting Charlotte Danielson as saying, "If all you do is judge teachers by test results, it doesn't tell you what you should do differently" (p. 244). If the goal of evaluations truly is professional development of educators, then suggestions and support for improvement need to be imbedded in the evaluation system.

Finally, Ravitch (2013) echoes this sentiment by turning accountability back to policymakers at the state and local level. She calls on these leaders to support schools, saying, "If they don't know how to help them, they should not be in charge. Accountability begins at the top, not the bottom" (p. 273).

Role of administrator in evaluation. Much has already been written about the potentially conflicting nature of the principal as judge *and* coach during the evaluation process. To reconcile this, Scudella (2015) writes about the need for "effective supervision," consisting of conversations about teaching, and feedback that is continuous and leads to reflection (p. 216). Principals need to be observant and understanding of the process of teaching, as well as set clear goals and improvement plans for individual teachers (Scudella, 2015, p. 217). According to Scudella, teachers want their administrators to provide them with open communication, frequent feedback, and directed goal setting during the evaluation process. They also want to have an opportunity to build a relationship with their administrator (p. 218).

Burton, Carper, and Wilburn (2011) agree that open communication from principals that provides "authoritative reassurance, thoughtful questioning, careful listening, and reflective practice" helps improve relationships between teachers and principals, and leads to better student achievement (p. 25). A key task for principals is to create a vision for learning and communicate it clearly to the staff, students, and community, provide a plan for achieving the vision, and give support along the way (Burton, Carper, & Wilburn, 2011, p. 29). According to Burton, Carper, and Wilburn (2011), the support comes in the form of valuing teacher perceptions during the evaluation process, and encouraging teachers to reflect on their teaching and impact (p. 30). Achievement scores should only be used to improve instruction, with the goal of evaluation being reflection and questioning of student data (Burton, Carper, &Wilburn, 2011, p. 30).

Finally, Ravitch (2013) recommends that principals be experienced, master teachers themselves (p. 131). This is essential if they are going to be observing teachers regularly and giving meaningful feedback that helps improve their practice.

Engaging teachers in evaluation reform. Another way to improve teacher perceptions of evaluation is to involve educators themselves in the process. Behrstock-Sherratt, Rizzolo, and Laine (2013) identify two ways to engage teachers in their evaluations (p. 57). First, district and school leaders should ask for teacher input on the design of evaluation systems. It is important that they genuinely want teacher feedback regarding evaluation systems; if they do, it will help gain teacher trust and buy-in (Behrstock-Sherratt et al., 2013 p. 64). According to Behrstock-Sherratt et al. (2013), teachers should also engage in self-evaluation, peer-evaluation, and principal evaluation (p. 58). This holistic approach also increases teacher commitment to the evaluation process.

Peer review. Peer review not only can serve as a way of engaging teachers in the evaluation process, but it also benefits other stakeholders. According to Goldstein (2014), having respected teachers review, observe, coach, and evaluate peers helps with the administrative burden for principals and helps gain the support of teachers in the evaluation system (p. 238). As previously stated, a common teacher complaint of the evaluation process is that principals cannot and do not give useful feedback because they do not understand the curriculum or the context of teaching. When master teachers serve as peer reviewers for a district, they focus all of their time and energy on providing specific coaching to struggling teachers, with the aim of improving their skills (Goldstein, 2014, p. 243). According to Goldstein (2014), this not only rewards the best teachers to

serve as mentors, but also saves districts money: the cost of replacing a dismissed teacher is $10,000, while peer review only costs $4-7,000 per teacher (p. 243).

According to Ravitch (2013) the Peer Assistance and Review (PAR) program in Montgomery County, Maryland is a prime example of how to implement the peer review evaluation process (p. 271). There, mentor teachers coach struggling teachers by helping them plan lessons, reviewing student work, and modeling effective teaching methods. They serve in this role for three years, and then return to the classroom (Ravitch, 2013, p. 271).

Goldstein (2014) reports that peer review is often criticized as being a "union ploy" that does not take evaluation seriously and only distracts from "actual" accountability measures (p. 240). However, according to Ravitch (2013), in Montgomery County, 200 teachers were dismissed under PAR, whereas in the decade before the program, only five teachers were fired (p. 271). Aside from those figures, peer review supporters say the number of dismissals is not the point, because peer review is intended to provide coaching to help an ineffective teacher become effective (Goldstein, 2014, p. 241). The system succeeds if a teacher is helped by PAR.

Conclusion

Teacher evaluation systems in the United States continue to evolve, and likely will keep doing so. Researchers will continue to ask questions about how effective these systems are at reaching their stated goals, as I will do in my own study of teacher perceptions of the Educator Effectiveness evaluation system in my state. The teachers who are evaluated under these systems are crucial sources to answer those questions, and hopefully their perceptions continue to shape education policy in our country.

Conclusion

My research question considers teacher perceptions of the Educator Effectiveness system of evaluation. This chapter has laid the groundwork by providing important information related to my research question. In the first section, I explored the history of the accountability movement in education, and where the current push for evaluation reform came from. In the second section, I examined different issues to consider when devising an effective evaluation system, taking into account differing perspectives, and attempts to reconcile them. In the final section, I considered other research into teacher and principal perceptions of evaluation systems throughout the country in order to provide a means of comparison for my own study.

The next section will provide an explanation of the methodology I plan to use for my research study to gauge teacher perceptions of the Educator Effectiveness system of evaluation. I will describe my research paradigm and method, the setting of my study, and my participants. I will outline the timeline, procedures, and tools I will use to conduct my study. Finally, I will explain how this study adds to the broader conversation around teacher evaluation within my district and state.

CHAPTER THREE

Project Description

My research question asks: What are teacher perceptions of the Educator

Effectiveness system of evaluation? To explore this, I used my Capstone Project to make

plans and tools for a hypothetical study that gauges teacher opinions of the evaluation

system using a mixed-method approach in order to triangulate the data I collect. This

chapter will provide an in-depth description of my project.

This research seeks to probe the opinions and perceptions of teachers regarding

their experiences with teacher evaluation. In the first section of this chapter, I will define

the research paradigm and the method I will use to complete this study. In the next

section, I will set the context for my project, the setting of the project, and the potential

participants. I will then describe the procedures and tools I will to use to collect data.

Finally, I will explain the timeline for my project and the audience with whom I hope to

share it. I will also discuss how this project adds to the conversation around the broader

topic of educator evaluation.

Implemented in 2014, Educator Effectiveness is still a new evaluation system in

my state, and as such, there is currently not a great deal of research into its impact on

teaching and learning in the state. The rationale for this study is to add to the body of

literature by exploring if teachers believe the system is improving teaching practice and

student learning.

Approach to Research

Research Paradigm

Study. To answer my question regarding teacher perceptions of the Educator Effectiveness evaluation system, I will use a mixed methods research paradigm, incorporating both quantitative and qualitative data. According to Cresswell (2014), "The 'mixing' or blending of data, it can be argued, provides a stronger understanding of the problem or question than either by itself" (p. 215). Using both methods will help me investigate the research question from different angles.

I want to compare survey data from a larger group of teachers to the answers I get during focus group interviews. My literature review showed some discrepancies in the data collected from these two sources in other similar studies, and I would like to see if the same is true for my research question. Having both quantitative and qualitative data will hopefully provide a broader picture of teacher perceptions of Educator Effectiveness in my school.

Presentation. Before collecting data, I will present a PowerPoint to potential participants at the school to give them background and rationale for my study. During this presentation to the teaching staff, I will use the principles laid out in Knowles' Andragogical Model to share my information with my adult audience (Knowles, Holton, & Swanson, 2005, p. 58). According to Knowles et al. (2005), andragogy, as opposed to pedagogy, focuses on adult learners and is based on six assumptions: need to know, self-concept, learner experience, readiness to learn, orientation to learning, and motivation (pp. 64-68).

The first assumption, that adults need to know *why* they should bother learning something, will help me make the relevancy behind my presentation and study clear to

my audience. Adult learners also have a strong self-concept, and do not enjoy being told

what to do or how to think, so I will strive to make my presentation as non-biased as

possible, so participants feel free to formulate their own opinions on the topic.

As far as learner experience goes, I must recognize that my audience will be filled

with educators with many different backgrounds, goals, motivations, interests, and

learning style, so individualization is important (Knowles, et al., 2005, p. 66). Almost

every teacher will have had experience with the Educator Effectiveness model of

evaluation, so it is pivotal that I make it clear that I want to tap into their individual

knowledge and experience with the system, and that I value their opinions.

The readiness to learn does not really apply to my presentation, since my audience

will already be familiar with my topic, but their orientation to learning is significant to

consider. Adult learners are life-centered, and want to know how their learning will help

them do things or deal with challenges in real life (Knowles, et al., 2005, p. 67). It will

be my task to apply my study to these teachers real-life experience with teacher

evaluations, emphasizing how my goal is to see what their perceptions of the

effectiveness of the system are.

The final assumption of andragogy is key to my presentation: motivation. I hope

to motivate teachers to take my survey and agree to participate in focus group

conversations after my presentation is complete, so it is imperative that they are

motivated to share their opinions with me. According to Knowles, et al. (2005), both

external and internal motivation can drive learning and decision-making, and that internal

factors are actually the most influential (p. 68). This is good news for me, for I will not

be offering any external rewards, such as money, gifts, etc., but will do my best to appeal

to my audience's internal motivation to share their opinions about Educator Effectiveness to help improve the system of evaluation in our school, District, and beyond.

Research Method

When I conduct my study, I will use a convergent mixed methods design. Cresswell (2014) writes, "The key assumption of this approach is that both quantitative and qualitative data provide different types of information – often detailed views of participants qualitatively and scores on instruments quantitatively – and together they yield results that should be the same" (p. 219). As indicated above, I'd like to see if the data is the same, or if there are some discrepancies between what teachers indicate on a survey, and what they say during an interview.

Qualitative method. My qualitative method will be focus group conversations. According to Mills (2014) a focus group is a "group interview" in which the researcher tries to "collect shared understanding from several individuals as well as to get views from specific people" (pp. 91-92). I will speak to at least two different focus groups, divided by age and teaching experience, to try to come to that collective understanding, and also get at individual opinions of teacher evaluation.

Quantitative method. My quantitative method will be an attitude scale survey. According to Cresswell (2014), this survey will allow me to provide a "numeric description of trends, attitudes, or opinions of a population by studying a sample of that population" (p. 155). I will use a Likert scale, as described in Mills (2014), to see if teachers agree, somewhat agree, somewhat disagree, or disagree with a series of statements about Educator Effectiveness (p. 102).

<div align="center">

Research Context

</div>

Setting

  Location. The setting of my project is an elementary school in a mid-sized city located in the upper Midwest of the United States. The school district has an approximate enrollment of 11,300 students, and is comprised of two high schools, three middle schools, 12 elementary schools, two charter schools, and one early learning center. The elementary school at which this study will place is located in the heart of the city, in a neighborhood that is considerably less affluent than others.

  Student demographics. The enrollment of the school is 271 students. Twenty-one percent of the students are disabled, and 74% are economically disadvantaged. The racial demographics for students are the following: 63.1% of students are white, 15.5% are Asian, 5.2% are Hispanic or Latino, 4.8% are Black or African-American, 2.6% are American Indian, 1.1% are Native Hawaiian or Pacific Islander, and 7.7% are classified by two or more races.

  Staff demographics. The teaching staff consists of 30 professionals: one principal, fourteen general education teachers, three special education teachers, four specialists (music, art, physical education, library), two speech teachers, a librarian, and five academic interventionists. All of these individuals are evaluated under Educator Effectiveness, and most, if not all of them have gone through a summary year in which they have been rated by the principal. These educators are qualified to comment on how they see Educator Effectiveness impacting teaching and learning at a school, district, and state level.

  There are also several pupil services staff members employed at the school who are not subject to evaluation under Educator Effectiveness, including a school counselor,

social worker, school nurse, and school psychologist. In addition, a number of para-professionals are also employed at the school to support the teachers and students, but they are also not subject to Educator Effectiveness.

In full disclosure, I have been teaching at this school for the past four years. Most of these staff members have been my colleagues throughout those four years. We have worked closely together in some aspect or another on a variety of projects, so I potentially bring a certain measure of bias with me to the study.

Participants

Survey. In the fall, I will invite all 30 certified staff members evaluated under Educator Effectiveness to participate in the study by means of a survey that will be emailed to them. The staff is at various stages in their teaching career, as well as various stages of the Educator Effectiveness process. The survey will account for this variability, asking specific questions regarding years of experience and whether they have had a summative evaluation year under the Educator Effectiveness model.

Focus groups. I will also conduct two focus group sessions, consisting of ideally 4 teachers in each group. These two groups will represent different age and experience levels. The teachers in the first group will have 10-plus years of teaching experience, while the second group of teachers will have less than 10 years teaching experience.

Procedures

First, I drafted a letter to the staff explaining the purpose of and background for my study (see Appendix A). I will present this information at a professional development meeting, handing out copies of the letter to the 30 certified staff members evaluated under Educator Effectiveness. I will also give my colleagues a paper copy of the survey

to take with them after the presentation, and ask that they all complete it within a week and return it to my mailbox (see Appendix B). I will assure them that their responses are anonymous.

Staff survey. After the meeting, I will send an email out to the teaching staff. The email will include key points from the presentation, a copy of the letter I presented at the meeting, and another copy of the survey itself, in case they misplaced the original. I will see how many people initially submit the survey, and if necessary, send out a follow-up email after a week has passed, requesting again that my colleagues take the time to fill it out.

Organizing focus groups. I will also privately approach teachers I have identified from each age group mentioned above, and ask if they would be willing to participate in a focus group conversation about Educator Effectiveness. I will remind them that their responses will be anonymous. Once I have received agreement from four members for each group, I will arrange a time and a place to meet off-campus that accommodates all participants' schedules.

I will then facilitate the conversations with each group, recording the audio to later go back and transcribe. I will use a prepared list of eight questions to start the conversation, but also will allow the conversation to develop organically, asking other questions that come up (see Appendix C). I will also tell the participants to feel free to follow up with me individually after the focus group session, if they have anything else they want to add that they thought of later, or that they did not feel comfortable sharing in the larger conversation.

Tools

I will use an Excel document to gather and analyze the data from the quantitative survey. For the qualitative focus group interviews, I will record, transcribe, and analyze them.

Staff survey. The survey questions consist of Likert scale attitude statements to see if/how strongly teachers agree or disagree with a series of statements. For example, "The Educator Effectiveness evaluation process enables me to grow as an educator" and "The Educator Effectiveness system is well-organized and easy to use" are statements that teachers will be asked to react to, using the Likert scale I provide.

Focus group questions. I will ask the same questions to both focus groups I conduct. As stated above, the conversations will be recorded for later transcription and analysis. The questions for the conversation will be much more open-ended, such as:

-In what ways is EE helpful for your practice? Do you believe it hinders it in any way?

-Do you feel the new model is more or less effective than other evaluation systems you have participated in?

-What do you perceive the purpose of EE to be for our school district and the state?

-Do you have any concerns with how the evaluation model is currently used, or could be applied in the future?

Data Analysis

Survey analysis. I will analyze the survey by organizing the responses into an Excel spreadsheet. I will look at the responses to see which Likert scale selection was chosen the most/least for each question.

Interview coding. For the focus group interviews, I will transcribe and code the data for common themes that come up from the participants.

<div align="center">

Implementation
</div>

Timeline

My intention is not to complete the action research component of this project during the Capstone process. Rather, it is a hypothetical research design to be implemented during the upcoming 2017-2018 school year. I worked to develop the research tools over the summer of 2017. I created a PowerPoint presentation that I plan to share with the certified staff in the fall, describing my interest in the topic, the background information, and sharing the methods I will use (see Appendix D).

When the school year resumes in the fall, teachers will begin a fresh cycle of Educator Effectiveness, and will be more prepared to reflect on their experiences with the evaluation system. I will present my PowerPoint, send out my survey, and conduct my focus group conversations within the first few months of school. I will then analyze the quantitative and qualitative data I collect.

Audience

My hope is that the research results will be meaningful to the participants. After I collect and analyze my data, I intend to follow up with staff and present another PowerPoint presentation about the data I collect from the survey and interviews at one of our weekly professional development meetings. Potentially, this information might be of interest to the school district, or even administrators at the state-level, and I would be happy to present and share it with them, as well.

Potential Extensions

School, district, and state level. Depending on the results of the study, there are several potential extensions for my project. On a building level, this research may start a conversation among the teaching staff about improvements that could be made within the school regarding educator evaluation that could lead to better teaching and learning. A presentation to the school board may be in order, as well, to encourage other schools in the district to survey their staff about their perceptions of Educator Evaluation. The final report of my data could also inspire the state Department of Public Instruction to consider possible modifications to the evaluation system, or at least recognize the need for further research.

Doctoral study. On a personal level, I may want to expand this project for a future doctoral thesis. It would be interesting to survey teachers at other schools within the district, especially with an aim to compare perceptions between teachers at the elementary and secondary levels. It might also be an opportunity to compare principal perceptions of evaluations to teacher perceptions, to see if there are any overlaps or discrepancies.

Impact

Growing the literature. This study will add to the very minimal literature on the topic of the Educator Effectiveness system of evaluation in my state. Since it is such a new system (it's been in statewide effect for only three years), there has not yet been much research on its implementation, efficacy, and reception by teachers. My research will give a look into one school in the state, and how teachers there believe this new evaluation model impacts their practice.

Continued reform. According to its mission statement, the Department of Public Instruction (DPI) aims to make "every child a graduate, college and career ready" in the state (Every Child a Graduate). One key component of this mission that the DPI has identified is a "fair and meaningful evaluation process" for teachers (Every Child a Graduate). Whether or not teachers actually perceive this process to be fair and meaningful is something this study aims to determine. If teachers are not invested in their development through the evaluation process, this may be an indicator that further reform is needed in this area. The DPI would be wise to seek feedback from its educators to find out what is working in the Educator Effectiveness evaluation system, and what could be improved. Hopefully this study will be able to identify both and give recommendations moving forward.

Conclusion

In this chapter, I have outlined the major components of my research methodology. I have identified my mixed-method research paradigm, and explained the quantitative survey and qualitative interview methods I will be using in order to triangulate my data. I have also described the school setting, teacher participants, and procedures for my study in detail. I discussed the tools I will use to gather data, and how I plan to analyze it. Finally, I presented important information regarding the implementation of my study, including the timeline, audience, and impact of the project.

It is important to again note that this research will be done in the fall after I developed the tools needed to complete my study. I also created a presentation for my Capstone class, which I plan share with my colleagues in advance of my study, informing them of the rationale, background, and procedures of my research.

In the next chapter, I will share what I learned from my project. I will also revisit the literature review and identify some key sources that informed my study. I will share the implications and limitations of my research, as well as some ideas for future research on the topic of educator evaluation.

CHAPTER FOUR

Conclusions

In this project I ask the following question: What are teacher perceptions of the Educator Effectiveness system of evaluation? I have spent the first three chapters explaining my connection to and interest in this question, reviewing the relevant literature to the question, and describing my project that seeks to answer it.

In this chapter I will share what I have learned throughout the Capstone process. I will also revisit the literature review, identifying key resources that shaped my project development. I will then explain the implications and limitations of the project, as well as the direction for future research. Finally, I will share how I plan to communicate the results of my project.

What I Learned

Throughout the course of this project, I have grown as a researcher and a writer. While pursuing my undergraduate degree, I worked on several faculty/student collaborative research projects, but this was the first time I embarked on such a large undertaking on my own. I'm grateful I had that undergraduate research background to prepare me for my work on my Capstone, but this opportunity to work independently provided new challenges and opportunities to develop my researching and writing skills. This experience also made me consider expanding my professional experience into other areas of leadership outside of teaching.

Research

I thoroughly enjoyed searching for resources applicable to my topic through the Bush Memorial Library. The process was user friendly and extremely helpful. It's hard for me to believe now that at one time this aspect of the project was the most daunting for me. I was afraid there would not be enough materials or that I would not have access to them. After only a week or so of using the online search engines provided by the university's academic library, I felt like I had a solid grasp of the process, and ended up locating more resources than I ever would have thought possible.

Another challenging aspect of researching this project was digesting and synthesizing all of the relevant information from the resources I found. The sheer amount of material was at first overwhelming, but I quickly developed a reading and note-taking system that worked for me to get through the books and articles in an efficient manner. The next step was organizing my notes into themes and sections, which is always enjoyable for me. I like to put things together. I relish the process of making connections and synthesizing information from disparate sources into one coherent whole. Thankfully, this complicated project offered me a great opportunity to utilize and hone this skill.

Additionally, I learned how rewarding it is to become somewhat of an expert on a specific topic. After reading, pondering, and organizing everything I researched about accountability reform and educator evaluation, I feel like I can speak and write knowledgably on the topic. I intentionally chose to pursue a question that is timely and important to my school, district, and state, so that when personal conversations and policy issues arise, I can feel confident in presenting an informed opinion, based on what I learned throughout this Capstone process.

Writing

Writing has always been an area of strength for me. As I said above, I enjoy synthesizing and communicating my thoughts and learning. This project gave me ample opportunities to do just that, but I really had to work hard to keep my writing strictly academic and avoid literary embellishments. My peer reviewer was very helpful in spotting occasions where I used clichés or colloquialisms, and I'm grateful for her keen eye and experience with academic writing.

I also found it challenging to describe in detail the steps I plan to take for my action research. There were several aspects of my research that I might have taken for granted as obvious and failed to include, but the Capstone Workbook kept me on track in thoroughly describing each element of my project. I recall my instructor telling our class that after reading the third chapter of our Capstone, readers should be able to go out and actually do the research themselves. That level of specificity was difficult for me at first, but ultimately I believe it has led to a strong project description.

Finally, I learned how to write and prepare a professional presentation. I've given several PowerPoint presentations in academic classes over the year, but I found it was a bit different process crafting the presentation for my colleagues at my school. Since I work with these people every day and have a great deal of respect for them, I wanted to make sure that my presentation was engaging and valued their own expertise and opinions – and would convince as many of them as possible to participate in my study. Since I typically teach children, it was new for me to consider how best to reach adult colleagues. I was grateful for the guidance of Knowles' work on adult learning to create the best presentation I could to answer my research question.

Potential Professional Goals

My experience with this project has also made me realize that I desire more opportunities for learning and leadership. While I'm not entirely sure what form that will take yet – academia, administration, or politics – I know that I am not finished thinking about and engaging in issues that affect education in my city, state, and nation. I will spend some time considering what next, exciting step to take in my career that will hopefully positively impact student learning.

Conclusion

I began this Capstone process intending to learn about my chosen topic, but I was surprised by how much my research and writing skills developed throughout the course of my studies. I feel more confident searching for academic resources, synthesizing information, writing an academic paper, and creating a professional presentation. I did not realize that I had as much room to grow in these areas, but I'm so glad I had the opportunity to improve these valuable abilities. This experience makes me want to find opportunities to apply them more.

In this first section I explained what I learned throughout the Capstone research process. In the next section I will return to a specific aspect of my research – the literature review – to consider which parts were the most important to my Capstone.

Revisiting the Literature Review

The literature review chapter of the Capstone was the most intimidating aspect of the process for me, but it ended up being the most rewarding, and the most helpful. By locating, reading, and synthesizing the relevant information related to my research question, it enabled me to craft my own project with a thorough understanding of what

has come before it. There are several resources that stand out as being particularly useful to me in the development of my Capstone.

Ravitch and Goldstein

Diane Ravitch is a well-known thinker, writer, and policy analyst in the field of education. She researches and writes about important educational issues and policies, as well as their impact on teachers and students. Her book *The Death and Life of the Great American School System* was quite influential to me when I first started my teaching career. I was interested to know what Ravitch had written on the topic of educator evaluation to help inform my Capstone.

This brought me to her most recent book for my research. Ravitch (2013) provided valuable context for the historical look at accountability reform in general, as well as specifics on teacher evaluation reform. In her thorough way, Ravitch (2013) lays out the policy decisions that led to the current education climate in the United States, and help me gain a deep understanding of the factors at play and the goals of the stakeholders involved.

Journalist Dana Goldstein wrote an equally detailed account of the history of the teaching profession and all of the political issues that surround it. Goldstein (2014) provided me with a useful background context to my research question, and offered valuable information on all of the important policy leading up to current education reform landscape in this country.

Danielson

Charlotte Danielson is one of the premier names in teacher evaluation. Since my state utilizes her *Framework for Teaching* in its educator evaluation system, I have had

the opportunity to learn about her approach and methods in great depth, and this information was extremely valuable to my project.

Danielson and McGreal (2000) was an important resource to help explain the goals of educator evaluation, as well as the difficulty in reconciling the professional development aspect of evaluation with the judgmental nature of it. Danielson (2007) is a comprehensive explanation of her methods, and is what the Educator Effectiveness evaluation system is based upon. Both of these resources were instrumental to my project, helping me frame the debate around teacher evaluation, and the way Educator Effectiveness addresses it. When I began this project, I knew that Charlotte Danielson would be a key part of my research, and I was grateful for the wealth of material and accessibility of the information.

Downing

The final part of my literature review that was quite helpful to my project is Downing (2016). This is a dissertation on teacher evaluation in Ohio schools, and how it correlates with job satisfaction. Despite the fact that this is a different research question than my own project, I was struck by one of the researcher's findings that helped influence my own research design.

Downing (2016) found that there was no relationship between the teacher evaluations and job satisfaction from her quantitative survey (p. 97). However, the open-ended comment section of her instrument revealed some negative perceptions of evaluations, with teachers sharing opinions about their dissatisfaction with the evaluation system (Downing, 2016, pp. 99-100). This discrepancy influenced my research design, leading me to pursue a mixed-methods approach.

When I first learned about the different research approaches, I was drawn to mixed-methods as being the most comprehensive. When it came time to start designing my own research, however, I was overwhelmed by designing both a quantitative and a qualitative instrument for my research, and was going to simply develop a survey. After reading Downing (2016), though, I was reminded of the importance of gathering both quantitative and qualitative data to help answer my research question, because as Cresswell (2014) says, "The 'mixing' or blending of data, it can be argued, provides a stronger understanding of the problem or question than either by itself" (p. 215). I'm very interested to see how the quantitative and qualitative data from my study reflects or contradicts each other.

Conclusion

While the literature review was the most difficult and time-consuming component of my Capstone, it was also very beneficial. It provided me with the knowledge base I needed to design my research, and craft a thorough, well-informed presentation for my colleagues on the issue of educator evaluation. I learned important things from everything I read, but I am particularly grateful for the sources I described above as being particularly influential.

In this section, I have revisited my literature review, identifying resources of particular importance and making connections to my project. In the next section, I will outline the implications for my project, including policy implications.

<center>**Implications**</center>

I chose my research question partially with the implications of my project in mind. I wanted to select a topic that was timely and relevant to myself and other teachers

in my school, district, and state. Teacher evaluation is something that affects all educators, and is intended to improve teaching and lead to better learning outcomes. My project puts teachers' opinions of the evaluation system in my school at the forefront, asking them whether or not they believe it is a viable, helpful model. As such, I believe that my research into Educator Effectiveness has several potential implications for the future of education in my state.

Giving Teachers a Voice

Under Educator Effectiveness, teachers have their performance measured at least every three years, more for novice teachers. They are rated by their evaluator's perception of their practice, based on a series of classroom observations and discrete teaching artifacts, and whether or not their students achieve the Student Learning Objective they set. My project is intended to give teachers a voice in whether or not they perceive the evaluation system is achieving its stated purpose of improving teaching and learning in the state. In essence, it is their opportunity to evaluate the evaluation system. It is my hope that they are honest in their responses, and use their voice to identify things that are working, and things that should be changed.

Another related implication is for teachers to brainstorm ways to improve the system. I hope that the focus group conversations lead to productive discussions and ideas to make the system potentially more viable for teachers. Educator evaluation is here to stay, but the debate over the best approach to identify and promote excellent teaching will likely be ongoing. I believe that teachers, who are directly impacted by evaluation, should have a place in that dialogue.

Policy Implications

Depending on the results of my study, there is potential for the data to be communicated to a broader audience. After my study is completed, I plan on creating another presentation to share the results with the teaching staff at my school. If there is interest, I would be willing to present the findings to stakeholders in education within my school district, or even at the state level. It is my hope that the data from my research could be used to add to other information assessing the effectiveness of the state's evaluation system, and potentially lead to reforms to make it even more beneficial and sustainable for teachers.

Personally, I believe that there is much that is good about Educator Effectiveness, but there are some aspects that I don't believe policymakers considered from a teaching standpoint that could be tweaked to make the evaluation system less burdensome and more supportive. It remains to be seen if the data from my study reflects that opinion, but at any rate, I sincerely hope that policymakers would be willing to hear teachers' perspectives of the evaluation system and make adjustments to better suit their needs.

Conclusion

My project will enable teachers to express their opinions on Educator Effectiveness, as well as ideas and concerns for the future of teacher evaluation in our state. Hopefully policymakers will take notice of the results of this and other similar studies, and use the findings to inform future reforms.

While this section focused on the potential implications of my project, the next section will outline the limitations.

<div align="center">

Limitations

</div>

My project was impacted by a few limitations. The first is the small sample size. I intentionally wanted to keep the scope of my Capstone manageable, and chose to limit it to my own elementary school. The small teaching staff of only 30 is not ideal, but it was the most practical sample for me to pull from. My hope is that most teachers choose to participate. Through my presentation, I will do my best to encourage each to do so.

The other limitation is the potential for bias in this project. The study will be conducted at the school where I am a teacher, so there remains a potential that I could inadvertently influence my participants' responses. The teachers themselves bring their own political and personal bias to this research, which will no doubt impact how they respond to the survey and focus group questions.

Finally, every teacher has a different experience with evaluation based on several uncontrollable variables. Teachers and evaluators are human beings, and therefore their practice and perceptions are not entirely reliable. Attitudes may change on any given day. This project will attempt to identify trends in opinions, but the unreliability of such responses is definitely a limitation of my study.

This section looked at the limitations of my project. The small sample size, the possibility for bias, and the unreliability of perceptions are all challenges to this study. The next section considers possible future research, some of which may overcome some of these very limitations.

Future Research

There are some obvious logical extensions to my project that would further enhance the research question. The first is to expand the study to other elementary schools within my district to achieve a larger sample size. It would also be interesting to

compare data from elementary teachers with middle and high school teachers, to see how their experiences with teacher evaluation compare.

Ideally, this research would expand beyond my own school district to consider the state at large. I think it is important to involve as many teacher voices in sharing their opinions of Educator Effectiveness. It may also be useful to compare principal perceptions of the evaluation system with teachers' perceptions. As evaluators, principals have an important perspective to add to the conversation of educator evaluation in the state.

Conclusion

In this chapter, I reflected on my experience creating my Capstone project. I shared the important things I learned, including honing my research and writing skills. I also revisited my literature review to identify the sources that best guided me on becoming familiar with the issues surrounding educator evaluation and designing an action research project to investigate teacher perceptions of Educator Effectiveness. I also recognized the implications and limitations of my research. The key implication of my project is giving teachers a voice in how they are being evaluated, but this is limited by the small sample size and the potential for researcher bias. Finally, I laid out ideas for future directions for research that would add to my own project.

This Capstone journey has been extremely rewarding for me, personally and professionally. I enjoyed selecting and learning about an issue of importance and interest to me, I look forward to conducting my research with my colleagues, for I believe they will find it relevant and thought provoking. I also am interested to see what direction teacher evaluation takes in the state and the nation over the next few years.

REFERENCES

2011 Senate Bill 461. *Wisconsin Act 166.* Wisconsin State Legislature, 2011. Retrieved

   March 11, 2017 (https://docs.legis.wisconsin.gov/2011/related/acts/166).

Behrstock-Sherratt, E., Rizzolo, A., & Laine, S. W. (2013). *Everyone at the table :*

   Engaging teachers in evaluation reform. Somerset, NJ: Jossey-Bass.

Bullis, B. (2014). *The perceived impact of teacher performance ratings on the teacher*

   evaluation process: Voices from the field. Retrieved from ProQuest Dissertations

   & Theses Global. (1550896503)

Burton, L., Carper, M.K., & Wilburn, V. (2011). The sociology of educational

   supervision and evaluation. *Journal of Cross-Disciplinary Perspectives in*

   Education, 4(1), 24-33.

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(1), 46-59.

Center on Education Policy. (2008). *Standards-based reform in the United States:*

   History, research, and future directions. (1st ed.). Washington, D.C.: Hamilton, L.

   S., Stecher, B. M., & Yuan, K.

Creswell, J. (2013). *Research design: Qualitative, quantitative and mixed method*

   approaches (4th edition). Thousand Oaks, CA: SAGE Publications.

Danielson, C. (2001). New trends in teacher evaluation. *Educational Leadership*, 58(5),

   12-15.

Danielson, C. (2007). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: ASCD.

Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: ASCD.

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

Department of Education. (1991). *America 2000: An education strategy – sourcebook*. Washington, D.C.

Dohms, E. (2015, October 20). EC schools survey: Staff satisfaction has declined. *The Leader-Telegram*. Retrieved from http://www.leadertelegram.com/News/Front-Page/2015/10/20/Survey-Staffsatisfactionhas-declined.html

Downing, P. R. (2016). *Effects of teacher evaluation on teacher job satisfaction in Ohio*. Retrived from ProQuest Dissertations & Theses Global. (10118324).

Educator effectiveness online training. (2016). Retrieved from https://dpi.wi.gov/ee/training-tools/training

Goals 2000: Educate America Act, P.L. 103-227, 103rd U.S. Congress (1994).

Goldstein, D. (2014). *The teacher wars: A history of America's most embattled profession.* New York, NY: Anchor.

Gottlieb, D. (2014). *Education reform and the concept of good teaching*. New York, NY: Routledge.

Jackson, S., & Remer, C. (2014). *Evaluating teachers: Opportunities and best practices. re:VISION*, No. 04, Part 2. Durham, NC.

Knowles, M. S., Holton III, E. F., & Swanson, R. A. (2005). *The adult learner* (6th

edition). London: Elsevier.

Luders, Bill. (2013, September 1). Wisconsin educators put to the test. *Wisconsin Watch*.

Retrieved from http://wisconsinwatch.org/2013/09/wisconsin-educators-put-to-

the-test/

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

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

Mills, G.E. (2014). *Action research. A guide for the teacher researcher* (5th edition).

Upper Saddle River, NJ: Pearson Education Inc.

Moss, J. T. (2015). *The Danielson model of teacher evaluation: Exploring teacher

perceptions concerning its value in shaping and improving instructional practice.*

Retrieved from ProQuest Dissertations & Theses Global. (1710482692)

National Commission on Excellence in Education. (1983). *A nation at risk: The

imperative for educational reform*. Washington, D.C. Retrieved from

https://www2.ed.gov/pubs/NatAtRisk/index.html

National Council on Education Standards and Testing. (1992). *Raising standards for

American education*. Washington, D.C.: Campell, Jr., C.A., Romer, R., Ambach,

G., Baker, E.L., Benzel, B.L., Bicouvaris, M., … Smith, M.

New Teacher Project. (2009). *The widget effect: Our national failure to acknowledge and

act on differences in teacher effectiveness*. (2nd ed.). New York, NY: Weisberg,

D., Sexton, S., Mulhern, J., & Keeling, D.

No Child Left Behind Act of 2001, P.L. 107-110, 107th U.S. Congress (2002).

Nuñez, I. (2015). Teacher bashing and teacher deskilling. In He, M.F., Schultz, B.D. & Schubert, W.H. *The SAGE Guide to Curriculum in Education*. (pp. 174-182). Thousand Oaks, CA: SAGE

Ravitch, D. (2013). *Reign of error: The hoax of the privatization movement and the danger to America's public schools*. New York, NY: Knopf.

Rogers, F. (1977). *Having an Operation*. Pittsburgh, PA: Family Communications, Inc.

Scudella, C. (2015). *Teachers' perceptions of administrative autonomy-supportive versus controlling behaviors: Factors that influence teacher intrinsic motivation and relationships with administrators*. Retrieved from ProQuest Dissertations & Theses Global. (1766589701)

Sheppard, J. (2013). *Perceptions of teachers and administrators regarding the evaluation process*. Retrieved from Georgia Southern University Digital Commons: Electronic Theses & Dissertations. (853)

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

U.S. Department of Education. (2009). *Race to the top program executive summary*. Washington, D.C. Retrieved from https://www2.ed.gov/programs/racetothetop/executive-summary.pdf

Wacha, M. (2016). *Teacher perceptions of the teacher evaluation process*. Retrieved from ProQuest Dissertations & Theses Global. (1352156360)

Wachter, B. (2014, August 31). Tests, educator evaluations will change in the Eau Claire school district. *The Leader-Telegram*. Retrieved from http://www.leadertelegram.com/News/Front-Page/2014/08/31/Tests-educator-evaluations-will-change-in-the-Eau-Claire-school-district.html

The educator effectiveness system user guide. (2014). Retrieved from

https://dpi.wi.gov/sites/default/files/imce/ee/pdf/teacherprocessmanual-draft.pdf

APPENDIX A

Participant Recruitment Letter

September 4, 2017

Dear Colleagues,

Last month I completed my Master's Degree at Hamline University. As part of my graduate work, I prepared for a research study into teacher perceptions of the Educator Effectiveness evaluation process that I hope to conduct during the month of September 2017. I am writing this letter to ask you to participate in my research. This will be the culmination of a great deal of work on my project, and I am hoping the results will be meaningful for our school, the District, and potentially the state. I may also publish or use the results of this study in a scholarly way in the future.

My research seeks to gauge the perceptions of teachers at our school about Educator Effectiveness and how you feel the evaluation system impacts your professional practice. I will use a confidential survey to gather information, as well as follow-up individual conversations. If you feel comfortable talking to me further about the topic of teacher evaluation, there will be a place for you to indicate this on the survey.

This research will offer little to no risk to you as a participant. All results will be confidential and anonymous. I will not use any identifying information for the school or district in any subsequent writings about the results. I will conduct follow-up conversations with the ease of the participant in mind.

Please remember that your participation in any part of this study is completely voluntary, and you may decline from participation at any time.

I have received approval to conduct this study from (*name omitted for confidentiality*) and (*name omitted for confidentiality*) Elementary School Principal. The results could potentially be included in a professional journal or presented at a professional conference. In any situation, your identity will remain confidential.

If you agree to participate, please fill out the survey I will hand out during my presentation on my research project. Again, remember that your answers to the questions will be completely anonymous and confidential. Please complete the survey as soon as possible. I'm asking that all surveys be returned to my mailbox by September 15, 2017.

If you have any questions, please feel free to contact me. Thank you so much for your participation.

Sincerely

Ms. Erin House
(*contact information omitted*)

APPENDIX B

Survey

	Agree	Somewhat Agree	Somewhat Disagree	Disagree
1. The Educator Effectiveness evaluation process enables me to grow as an educator.				
2. The Educator Effectiveness process is well organized and easy to use.				
3. I have enough time to complete all aspects of the Educator Effectiveness process.				
3. I have adequate support during the Educator Effectiveness process.				
4. Post-observation feedback from my supervisor helps me improve my teaching practice.				
5. Reflecting on my teaching during the Self-Review process helps me improve my practice.				
6. Student Learning Objective goals are a fair way of measuring student success.				
7. I am encouraged to set ambitious Student Learning Objectives.				
8. I worry about being penalized if I fail to meet my Student Learning Objective goal.				
9. I use the Danielson Framework for Teaching to identify my areas of improvement.				
10. I understand what my school district is planning to do with the results of the evaluation process.				
What grade level do you teach?	PreK-1	2-3	4-5	Multi-grade
How long have you been teaching?	PreK-1	2-3	4-5	Multi-grade
Would you be willing to answer some follow-up questions?	Yes	No	If yes, provide email below:	

Please return to Erin House's mailbox by September 15th, 2017. Thank you for your help!

APPENDIX C

Focus Group Questions

1. What do you believe is the purpose of the Educator Effectiveness evaluation system for our school district and the state?

2. In what ways is EE helpful for your practice? Do you believe it hinders it in any way?

3. Do you feel the EE model is more or less effective than other evaluation systems you have participated in?

4. What do you think is a more important aspect of EE: teacher development or quality assurance?

5. Do you think that the components of EE (artifacts, SLO, observations) provide a fair reflection of your classroom practice?

6. Do you feel encouraged/supported to set rigorous goals for your Student Learning Objectives? Why or why not?

7. Do you have any concerns with how the evaluation model is currently used, or could be applied in the future?

8. Do you feel like your professional practice has improved in the three years of EE? If so, how much of that do you think is a reflection of the EE evaluation process?

APPENDIX D

Presentation Slides
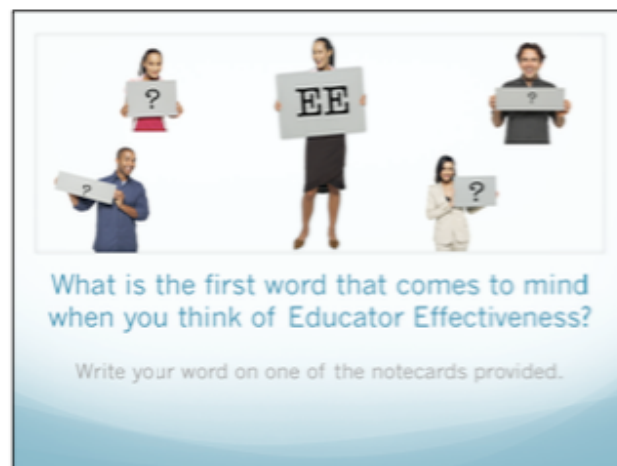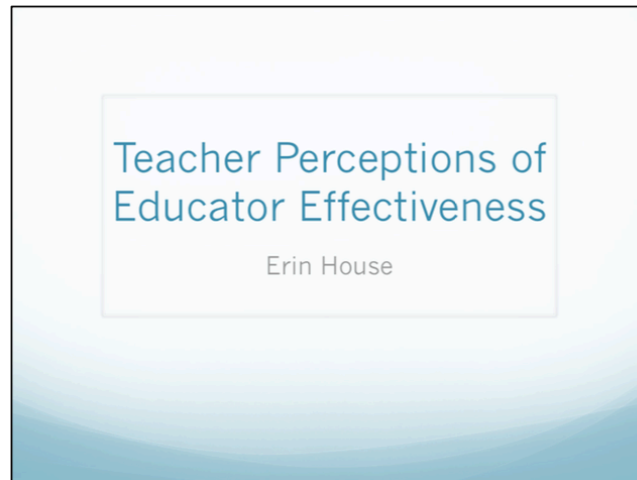
-To get started, I want you all to read the prompt on the screen and answer it on one of the notecards I've placed on your table. This does not have to be shared with anyone. In a minute or so, I'll ask if any volunteers want to share out their responses.
-(Wait for everyone to finish.)
-Does anyone feel comfortable sharing their word?

-I want to dig deeper into teacher perceptions of our evaluation system in the state.
-I'm hoping you all will help me answer my questions by participating in a confidential research project.

-Raise your hand if you ever completed a Master's thesis or capstone.
-This past year I completed my coursework to earn my Master's in Teaching at Hamline. The biggest part of that work was the Capstone project.
-We were advised to come up with topic that was meaningful to us that we wanted to dive into deeper.
-As I'll discuss further, teacher evaluation is a hot topic in ed. reform, and something we here in our state have been directly experiencing the last four years with EE.
-Also, I personally have been evaluated frequently as a TFA corps member, and as a new teacher to this district. Last year was the only time I wasn't evaluated.
-I have my own opinions, but I want to know what everyone else thinks, too, to see if there are trends.
-I think teacher evaluation is here to stay, but I personally hope it can be a useful tool for teacher's professional development.

Why do we as an educational community need to evaluate teachers and principals?

Please talk at your tables in small groups.

-Please take some time to discuss the question prompt in small groups.
-Think about your own experiences, and maybe some things you've heard from trainings about different evaluation systems you've experienced in your career.
-(Take volunteers who are willing to share what their group discussed.)

Two Approaches

- Formative evaluations = feedback for professional growth and improvement
  - Use observations to measure
  - Favored by educators
  - Reflects complex nature of teaching

- Summative evaluations = screen out "bad" teachers
  - Use student data to measure
  - Favored by politicians
  - Accountability to constituents

Source: Danielson & McGreal, 2000, p. 8.

-Evaluations can potentially serve two purposes: formative and summative.
-Think of these like the evaluations you give your students: one is focused on growth, the other on mastery.
-These two approaches are often at odds with each other when evaluating teaching ability and assigning ratings.
-Which one of these do you all favor?

Merging Both

- Danielson (2001) suggests a balance of formative and summative approaches when evaluating teachers.
  - Differentiation for novice/experienced teachers
  - Teachers take an active role in development
  - Focus on "professional conversations" to encourage reflection and mutual learning
  - Use student achievement data only as one of many measures, and ensure it is equitable and reliable.

Source: Danielson, 2001, pp. 14-15.

-Charlotte Danielson (whom most of you know well) seeks to merge both formative and summative approaches to evaluation through her approach and framework.
-She believes the following:
-Novice teachers need more frequent evaluations than experienced ones.
-Teachers should reflect and review themselves, identify areas of growth, set goals, and track progress.
-Conversations between colleagues and with principal/evaluator are central to development.
-Data is a part of the picture, and should drive conversations – teachers and evaluators should work backwards from data to hold coaching conversations; meant to enhance instruction, not merely judge it.

Educator Effectiveness

- Our state has chosen Danielson's approach when drafting its educator evaluation model.

- Discussion questions:
  - What are some parts of EE that reflect Danielson's philosophy of seeing the evaluation process as a time for reflection, review, and personal growth?
  - How is student achievement data utilized in EE?

-As you know, our state uses Charlotte Danielson in our evaluation model.
-Most of you have been trained in and evaluated under the Danielson framework. Considering your experience, discuss the question with your tablemates.
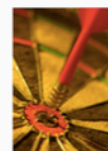-(After talk time, share.)

-(Read through questions, show video.)
-(See if anyone would be willing to share reactions with group.)

-As stated in the video we just watched, this is the goal of EE.
-(Ask volunteer to read it.)
-After three years of its implementation, I want to know: Is it working?

My Project

- Teacher perceptions of EE
    - What is helpful?
    - What could be improved upon?
    - Does it enhance or diminish the teaching profession?
    - What is the future of evaluation in the state?

-You've seen the training video, you've read the goal, now I want to know what you think.
-(Ask different people to read the questions.)
-This system is meant for YOUR growth and development, YOUR improvement, YOUR impact on kids being enhanced.
-Do you think it's a good system? Have you bought in? Is this a useful growth and development tool?
-If so, why? If not, why not?

My Project, Your Voice

- Letter
    - Explaining rationale, methodology, application

-I've created three tools to help me gather your perceptions of EE.
-The first is a letter, explaining my project and requesting your participation
-(Hand out letter)

-(Read through highlights of letter, especially confidentiality, approval, and deadlines.)
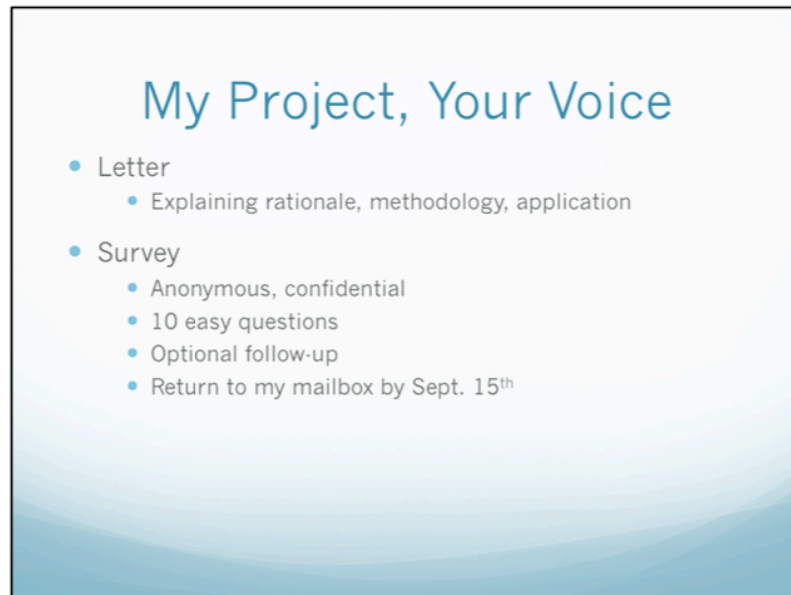
-The survey is something I would love everyone in this room to complete. The more participants, the better.
-(Go through list about survey, then hand it out.)

-(Go through questions)
-(Highlight the last three questions on the bottom)

-I will invite a lucky few of you to participate in the focus group conversations.
-I am ooking for 3-4 teachers from two groups of experience levels.
-The teachers in the first group will have 10-plus years of teaching experience, while the second group of teachers will have less than 10 years teaching experience.
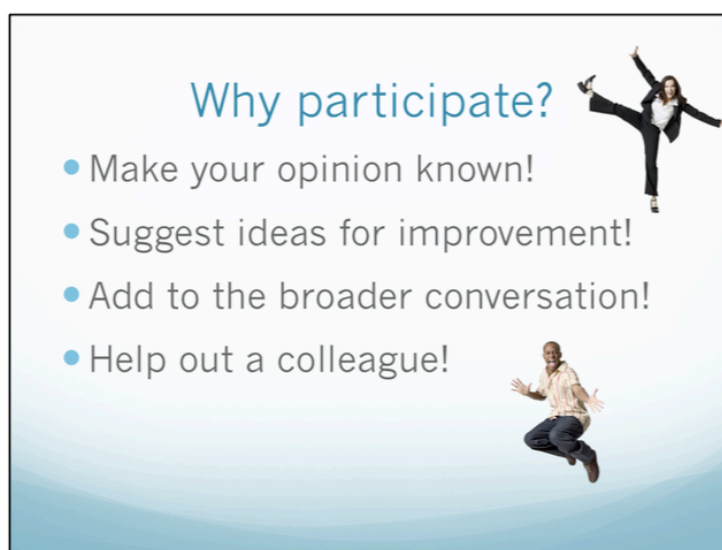
Focus Group Questions

1. What do you believe is the purpose of the Educator Effectiveness evaluation system for our school district and the state?

2. In what ways is EE helpful for your practice? Do you believe it hinders it in any way?

3. Do you feel the EE model is more or less effective than other evaluation systems you have participated in?

4. What do you think is a more important aspect of EE: teacher development or quality assurance?

5. Do you think that the components of EE (artifacts, SLO, observations) provide a fair reflection of your classroom practice?

6. Do you feel encouraged/supported to set rigorous goals for your Student Learning Objectives? Why or why not?

7. Do you have any concerns with how the evaluation model is currently used, or could be applied in the future?

8. Do you feel like your professional practice has improved in the three years of EE? If so, how much of that do you think is a reflection of the EE evaluation process?

-(Read over questions.)
-Look for a personal email invitation from me, and/or I will approach you about it in person.
-This is a great chance for us to talk in depth about your experiences with EE, and it would be really helpful for my project!

Why participate?

- Make your opinion known!
- Suggest ideas for improvement!
- Add to the broader conversation!
- Help out a colleague!

-It will be fun! It will be painless! It will be easy! It will be important! It will be so helpful!
-Educators are the people who are participating in Educator Effectiveness, and the system is meant to help educators, so let people know if it is working for you, or not.
-I've worked very hard on this project, but what makes it meaningful is if I get data to analyze and reflect on.
-I'm very passionate about and interested in this topic.
-I may expand this work into further graduate study or present it the results in a professional capacity.

Conclusion

- Teacher evaluation is here to stay.
  - Accountability reforms
  - Push to educate all students
- Future reforms could potentially be based on teacher input.
  - Changing political climate
  - More research
- Find what works to improve teaching and learning for our school, district, state, and country.

-We're going to have evaluations. As public school teachers, we are accountable to taxpayers. NCLB and Race to the Top created the push to ensure educational quality for all of the nation students, and puts that responsibility firmly on schools and individual educators.
-Who knows what direction teacher evaluation will go in the next few years. It is my hope that research studies and surveys such as mine will help determine that. Teachers are the ones directly participating in and affected by an evaluation system, so their perception of its usefulness is imperative in determining how effective the system is.
-All stakeholders in education agree that teaching and learning are of paramount importance. It's important to work together to advance both within our society

References

Danielson, C. (2001). New trends in teacher evaluation. *Educational Leadership*, 58(5), 12-15.

Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: ASCD.

Educator effectiveness online training. (2016). Retrieved from https://dpi.wi.gov/ee/training-tools/training

Note: All clipart in this presentation is copyright © Microsoft® PowerPoint® 2011.