12-2016

Teacher Perceptions of the Impact of the North Carolina Teacher Evaluation Process on Teacher Effectiveness, Professional Growth, and Attitudes Toward Teaching

Kim H. Case
Gardner-Webb University

Teacher Perceptions of the Impact of the North Carolina Teacher Evaluation Process on
Teacher Effectiveness, Professional Growth, and Attitudes toward Teaching

By
Kim H. Case

A Dissertation Submitted to
the Gardner-Webb University School of Education
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Education

Gardner-Webb University
2016

Approval Page

This dissertation was submitted by Kim H. Case under the direction of the persons listed below. It is submitted to the Gardner-Webb University School of Education and approved in partial fulfillment of the requirements for the degree of Doctor of Education at Gardner-Webb University.

_____   _____
Stephen Laws, Ed.D.   Date
Committee Chair

_____   _____
Jeffrey R. Church, Ed.D.   Date
Committee Member

_____   _____
Cheryl Lutz, Ed.D.   Date
Committee Member

_____   _____
Jeffrey Rogers, Ph.D.   Date
Dean of the Gayle Bolt Price School
of Graduate Studies

Abstract

Teacher Perceptions of the Impact of the North Carolina Teacher Evaluation Process on Teacher Effectiveness, Professional Growth and Attitudes toward Teaching. Case, Kim H., 2016: Dissertation, Gardner-Webb University, Teacher Evaluation/Teacher Effectiveness/Teacher Growth/Teacher Accountability/North Carolina Teacher Evaluation Process (NCTEP)/North Carolina Educator Evaluation System (NCEES)

This study's purpose was to examine teacher perceptions of the impact the North Carolina Teacher Evaluation Process (NCTEP) has on their professional growth, effectiveness, and attitudes toward teaching. Literature review reveals a longstanding interest in teacher evaluation for ensuring teacher quality and enhancing professional growth. Controversy exists in public and educational arenas as to how both purposes can be accomplished within a single evaluation system.

North Carolina evaluation reforms have mirrored nationwide efforts to use teacher evaluation as a tool for improving teacher effectiveness. The incorporation of value-added measures (VAMs) into teachers' summative ratings and implementation of the North Carolina Educator Evaluation System (NCEES) for online data collection were outcomes of North Carolina's Race to the Top grant.

In this study, teachers from a midsized school district in western North Carolina completed the Teacher Evaluation Profile (modified). This online survey identified correlations among key evaluation attributes and outcome ratings for overall quality of NCTEP, its impact on attitudes toward teaching, professional growth, and teacher effectiveness. Narrative responses were generated through survey comments and follow-up interviews.

Findings from this study supported the conclusions that teacher perceptions of the impact of NCTEP on professional growth, teacher effectiveness, and attitudes toward teaching differed based on years of teaching experience and grade level taught at the time of the last evaluation. Teacher attributes did not show significant relationships to outcome ratings.

Qualitative data indicated that teachers perceive that NCTEP does not include all aspects contributing to teacher performance. Teachers report that student factors such as motivation, attendance, behavior, and intellectual qualities should be taken into consideration in NCTEP.

Table of Contents

Chapter 1: Introduction

Nature of the Problem

Ensuring that students receive a quality education from a qualified teacher is a foundational goal for public schools (Gabriel & Allington, 2012). In even the earliest days of American public education, the acquisition and development of pedagogical skills was considered an important element of effective teaching, and instructional supervision focused on improving instruction (Marzano, Frontier, & Livingston, 2011); however, determining the characteristics and practices of an effective teacher and how best to measure those components have proven to be both complex (Darling-Hammond, Wise, & Pease, 1983) and controversial (Weisberg et al., 2009). While rating teacher performance is nothing new nor is it uncommon as a worldwide practice (Orphanos, 2014), growing political and public interest in measuring teacher effectiveness has evolved over the past several decades (Darling-Hammond, 2014) and has ignited changes in the development and structure of educator evaluation systems throughout the United States (Databases on State Teacher and Principal Evaluation Policies, n.d.).

The prologue of *Rethink, Rebuild, Rebound* states, "It has been said that education is simply the soul of a society as it passes from one generation to another" (Balls, Eury, & King, 2011, p. xii). As the literature reveals, an understanding of the passing of values from generation to generation is relevant to the study of teacher evaluation (Gabriel & Allington, 2012). Stakeholders in the field of teacher evaluation are numerous and range from students and teachers working in classrooms to politicians, corporations, and nonprofit organizations (Doyle & Han, 2012). Generally, teachers agree that evaluation should have two primary functions: measuring teacher effectiveness and enhancing teacher professional growth (Marzano, 2012); however, due to competing ideologies of

stakeholders, designing and implementing an evaluation system that leads to results in those arenas is elusive (Taylor & Tyler, 2012).

Notably, teacher quality has long been recognized as having a significant impact on student success (Marzano et al., 2011). For example, in the earliest days of public schools, interest in the development of effective instructional practices was evident. Marzano et al. (2011) explained,

> The period from the beginning of formal education in the United States up to the mid-1800s saw the dawning of the awareness that pedagogical skills are a necessary component of effective teaching. Although there was little or no formal discussion about the specifics of these skills, the acknowledgment of their importance might be considered the first step in the journey to a comprehensive approach to developing teacher expertise. (p. 13)

Over time, the processes of teacher supervision and evaluation grew in complexity as they were influenced by the views of emerging educational theorists (Marzano et al., 2011). Two conflicting views of education reform which impact the modern-day rationale that teacher evaluation should measure teacher effectiveness and enhance professional growth emerged from the philosophies of Frederick Taylor and John Dewey (Marzano et al., 2011). While Taylor focused on the use of scientific measurement of tasks as a means to improve worker productivity, Dewey proposed that the main responsibility of education was to foster a democratic society (Marzano et al., 2011). Taylor's theories supported efforts by Ellwood Cubberly, in 1916, to popularize the practice of measuring teacher performance by using a scientific approach to monitor teachers and reinforced William Wetzel's work in the 1920s that

> recommended three components as the basis for scientific supervision: the use of

aptitude tests to determine the ability level of each child; the establishment of

clear, measureable objectives for each course; and the use of reliable measures for

student learning. (Marzano et al., 2011, p. 15)

On the other hand, the influence of Dewey's democratic ideals can be seen in the

development of the clinical supervision model of the late 1960s in which teacher

professional growth is stimulated through reflective conversations between teachers and

observers (Marzano et al., 2011). Basic tenets of these theories are identifiable in

educator evaluation processes of today (Weisberg et al., 2009).

As predicted by the National Institute of Education (Wise, Darling-Hammond,

McLaughlin, & Bernstein, 1984), interest in improving evaluation has grown steadily

since the late 1970s and continues to receive national attention as teacher evaluations are

attached to high-stake outcomes (Harris, Ingle, & Rutledge, 2014). In response to

political and public outcries demanding that teachers prepare students to graduate from

public schools ready for success in college and careers, traditional methods of teacher

evaluation as it relates to teacher quality have been scrutinized (Weisberg et al., 2009).

Binary evaluation systems in which teachers receive a satisfactory/unsatisfactory rating

have been questioned due to the fact that these methods have produced a large number of

satisfactory rankings that could only minimally be linked to teaching and learning

(Weisberg et al., 2009). Systems that utilize checklists have also been criticized because

they do not lend themselves to meaningful conversations about instructional practices

between teachers and administrators (The Teaching Channel, 2013). Harris et al. (2014)

explained, "While formal teacher evaluation tools have been in practice for decades, they

give nearly all teachers the highest possible ratings and provide almost no information

about the technical or instructional core of teaching" (pp. 74-75). Jim Hull (2013), senior

policy analyst for the National School Boards Association's Center for Public Education, reinforced this opinion, stating, "For decades, teacher evaluations were little more than a bureaucratic exercise that failed to recognize either excellence or mediocrity in teaching" (p. 1).

Concerns about the quality of evaluation tools are only one aspect of the issues enveloping teacher evaluation (Toch, 2008). Weisberg et al. (2009) contended that the existence of an underlying cultural mindset in schools assumes that teachers are synonymous to interchangeable parts. According to these authors, both teacher expectations and observer perceptions of teacher evaluations are skewed due to this deep rooted attitude. They point to their research as an illustration of "policy framework that rarely considers teacher effectiveness for key decisions" and is descriptive of "schools [that] are indifferent to instructional effectiveness–except when it comes time to remove a teacher" (Weisberg et al., 2009, p. 4). Findings such as these have led to concerns about the ability of principals to evaluate teacher performance effectively (Harris et al., 2014) and the quality of the evaluation systems (Hull, 2013). Further criticism of teacher evaluation processes stems from the occurrence of student achievement gaps between some ethnic groups that have continued over several decades (Rooney et al., 2006). These gaps in student performance assume weaknesses in teacher effectiveness, prompting appeals for better tools to measure and improve teacher quality (Lee, 2011).

Marzano (2012) stipulated that the problems with current systems of evaluation exist because the two purposes of the evaluation process are at odds. Identifying these purposes as determining teacher effectiveness and enhancing professional growth, Marzano contended that "***Measuring*** teachers ***and developing teachers are different purposes*** with different implications" (Marzano, 2012, p. 16). He asserted that systems

for measuring teacher effectiveness will be inherently different for systems that are created to improve teacher performance. Darling-Hammond (2014) maintained that both objectives can be achieved through a single process if a comprehensive approach with multiple measures of evaluation is utilized. She did, however, question the appropriateness of educator evaluation systems that heavily rely on value-added measures (VAMs), a statistical method of measuring teacher quality, to determine teacher effectiveness (Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012). Nonetheless, as debates about teacher evaluation processes continue, trends in education indicate that many states are moving forward with the use of multiple measures, including VAMs, as components of new evaluation systems (Hull, 2013).

Purpose of the Study

The purpose of this study was to examine teacher perceptions of the impact the teacher evaluation process has on professional growth and teacher effectiveness as well as how the evaluation process impacts their attitudes toward teaching. While much research has been done to explore measures of teacher effectiveness (MET Project, 2013) and new evaluation systems are being implemented nationwide (Mead, 2012), less research is available about educator attitudes toward these new evaluation tools (Donaldson, 2012; Lee, 2011). In North Carolina, for instance, information about teacher perceptions of changes to the North Carolina Teacher Evaluation Process (NCTEP) stemming from Race to the Top (RttT) grant requirements has only recently been published (Marks, Fuller, Guthrie, Henry, & Stallings, 2015). For educators in the field, the outcome of teacher evaluations can have personal and professional implications (Sawchuk, 2015); thus, recognizing that teacher evaluation has the potential to enhance professional growth and teacher effectiveness (Stiggins & Duke, 1988), it is important to

understand how educators perceive the evaluation process as it relates to their teaching practice (Lee, 2011).

This study examined the perceptions of teachers in a midsized school district located in the foothills of North Carolina. Data that were gathered detailed their experiences with NCTEP during their most recently completed annual evaluation cycle. This study determined educator perceptions of NCTEP as it impacts their professional growth and teacher effectiveness as well as the impact the overall evaluation experience has on teacher attitudes toward teaching. Research data assisted in determining if and how NCTEP is accomplishing its intended purpose: "to assess the teacher's performance in relation to the North Carolina Professional Teaching Standards and to design a plan for professional growth" (Public Schools of North Carolina, 2015, p. 4).

The results of this study established whether teacher perceptions of the impact of the evaluation process on professional growth and teacher effectiveness align with the stated purposes of NCTEP (Public Schools of North Carolina, 2015). In addition, the study provided information about the impact the evaluation process has on teacher attitudes toward teaching. These results will help inform ongoing implementation planning and recommended staff development necessary for the successful use of NCTEP at both state and district levels. Additionally, results will identify attributes of the evaluation process that are in need of revision.

Background and Significance of the Problem

The complexities of teacher evaluation involve both how teacher performance is best measured (Darling-Hammond, 2014) as well as how ratings of teacher effectiveness should affect teachers and schools (Toch, 2008). The values that stakeholders possess with regard to education and the role of teachers have directed (Israel, 1978) and continue

to guide (Gabriel & Allington, 2012) the evolution of teacher evaluation (Weisberg et al., 2009). Understanding the development and current state of teacher evaluation processes in the United States requires an in-depth look at the public, political, and educational forces that have given it shape (Israel, 1978) and an examination of how advances in technology have altered the methods by which teacher effectiveness is measured (Betebenner et al., 2012).

Early calls for reform in teacher quality were made during the 1978 Chief Council of State School Officers (CCSSO) Summer Institute (Israel, 1978). At the gathering, discussions primarily focused on strategies intended to improve preservice teacher preparation but also suggested that in-service teachers could benefit from similar methods, including frequent observations with immediate feedback and ongoing professional development (Israel, 1978). As demands for greater teacher quality and accountability became more pronounced, teacher evaluation increased in importance which prompted the National Institute for Education to sponsor extensive study of teacher evaluation (Darling-Hammond et al., 1983). In one such study, researchers identified two essential components underpinning evaluation systems: (1) teaching effectiveness and (2) organizational and implementation theory (Darling-Hammond et al., 1983). Based on thorough literature review of these components, researchers hypothesized that "four minimal conditions for the successful operation of a teacher evaluation system" (Darling-Hammond et al., 1983, p. 320) were necessary. Darling-Hammond et al. (1983) described these four conditions as follows:

- All actors in the system have a shared understanding of the criteria and processes for teacher evaluation;

- All actors understand how these criteria and processes relate to the dominant symbols of the organization, that is, there is a shared sense that they capture the most important aspects of teaching, that the evaluation system is consonant with educational goals and conceptions of teaching work;

- Teachers perceive that the evaluation procedure enables and motivates them to improve their performance; and principals perceive that the procedure enables them to provide instructional leadership;

- All actors in the system perceive that the evaluation procedure allows them to strike a balance between control and autonomy, in describing this balance as "between adaptation and adaptability, between stability to handle present demands and flexibility to handle unanticipated demands" (Weick, 1982, p. 674); that is the procedure achieves a balance between control and autonomy for the various actors in the system. (p. 320)

More than 30 years later, teacher quality and accountability remain in the political and public spotlight (Shakman et al., 2012). Federal funding requirements for the RttT initiative and flexibility waivers in No Child Left Behind (NCLB) have fueled the emphasis on teacher effectiveness and led to the reform of educator evaluation processes (Darling-Hammond, 2014). Darling-Hammond (2014), seasoned researcher of teacher evaluation, pointed out, "virtually everyone agrees that U.S. teacher evaluation needs an overhaul" (p. 5), explaining that "problems are long-standing [and] . . . were obvious when my colleagues and I first studied U.S. teacher evaluation systems in the early 1980s" (p. 5).

Attempting to address these issues, an unprecendented number of legislative

actions surrounding state educator evaluation systems were enacted between 2010 and 2013 (Mead, 2012). Throughout the United States, individual states reacted to federal initiatives and other political demands by legislating the development and implementation of new educator evaluation systems (Mead, 2012). These performance-based systems utilize multiple measures of teacher performance such as observations, analysis of artifacts, peer and student reviews, self-reflections, and participation in professional development activities (Mead, 2012). Since mandates from the 2009 RttT competition required that teacher evaluation systems include multiple rating categories and student growth measures, many state evaluation processes were reformed to incorporate these components (Databases on State Teacher and Principal Evaluation Policies, n.d.). Under these guidelines, binary systems have become unacceptable and recently developed VAMs, which rely on standardized test scores to generate student growth rates, have become a new and often controversial element of educator evaluation systems in many states (Hull, 2013; Shakman et al., 2012).

     With the creation of new comprehensive evaluation systems, the stakes surrounding teacher evaluations have risen (Darling-Hammond, 2014; Harris et al., 2014). Discussions of how teacher effectiveness measures can be used to retain, dismiss, or compensate teachers are common (Taylor & Tyler, 2012; Weisberg et al., 2009). According to Lee (2011), the prevalent opinion driving reform efforts is that "if you have good teachers, you will have good student achievement" (p. 98). This common view recognizes that teachers are the main factor driving student achievement and has placed emphasis on measuring student achievement through VAMs (Lee, 2011).

     Proponents of VAMs exhort their ability to measure and predict performance in a quantifiable manner (SAS, 2015). Others, including teachers and seasoned researchers of

teacher evaluation, are critical of VAMs (Collins, 2014; Darling-Hammond, 2014;

Marzano, 2012). Nonetheless, numerous states are moving forward with the inclusion of

VMAs in their new teacher evaluation systems (Hull, 2013), even as research into the

effects of VAMs on teacher ratings is in flux (Davis, Bangert, Comperatore &

Smalenberger, 2015). As a result, some states are faced with litigation that challenges

various aspects of these new evaluation systems (Sawchuk, 2015).

In North Carolina, teacher evaluation has flowed with the political tide, as the

state participated and won an RttT grant in 2010 (NCLEG, 2010). Although revisions to

NCTEP began in 2008 (Public Schools of North Carolina, 2009), the RttT grant funding

enabled the acceleration of reforms to NCTEP (NCLEG, 2010). Multiple revisions of

NCTEP between 2008 and 2015 are indicative of these rapid changes (Public Schools of

North Carolina, 2009, 2012b, 2015).

The earliest implementation of NCTEP was approved in 2008 by the North

Carolina State Board of Education and reflected the revision of the Core Standards of the

Teaching Profession, which was adopted in 1997, to the North Carolina Professional

Teaching Standards (Public Schools of North Carolina, 2009). These teaching standards

formed the basis for not only teacher evaluation but for teacher preparation and

professional development as well (Public Schools of North Carolina, 2012b). Further

revisions to NCTEP that added abbreviated evaluations for teachers with career status

were made in September 2011 (Public Schools of North Carolina, 2015). In January

2015, modifications were again made to remove references to career status and

probationary teachers and to outline requirements of three new evaluation cycles:

comprehensive cycle, standard cycle, and abbreviated cycle (Public Schools of North

Carolina, 2015).

The 2015 NCTEP manual can be accessed online at the North Carolina Public Schools website (Public Schools of North Carolina, 2015). This updated manual identifies the purpose and goals of the evaluation process, defines the North Carolina Professional Teaching Standards, explains the eight components of NCTEP, and provides examples of evaluation rubrics and related forms. As reported on North Carolina Department of Public Instruction's (NCDPI) Educator Effectiveness Model webpage, the North Carolina Educator Evaluation System (NCEES), an electronic data collection system, is used during NCTEP (Public Schools of North Carolina, n.d.a). NCEES houses all evaluation information on a web-based platform hosted by True North Logic (Lynn et al., 2013).

The North Carolina Professional Teaching Standards serve as the basis for NCTEP and also underpin teacher preparation and professional development (Public Schools of North Carolina, 2013b). In Standards 1-5, the skills and knowledge that teachers are expected to possess are defined (Public Schools of North Carolina, 2015). Standard 6, which was added to NCTEP by the North Carolina State Board of Education in February 2012, stated, "The work of the teacher results in acceptable, measurable progress for students based on established performance expectations using appropriate data to demonstrate growth" (Public Schools of North Carolina, 2015, p. 12). Whereas the evaluation of Standards 1-5 is completed through the work of observers, Standard 6 is based on "student growth value as calculated by the statewide growth model for educator effectiveness" (Public Schools of North Carolina, 2015, p. 40). In North Carolina, the Education Value-Added Assessment System (EVAAS), a statistical analysis system, calculates student growth rates and determines teacher effectiveness that is included in Standard 6 (Lynn et al., 2013).

According to the NCTEP manual, "The intended purpose of the North Carolina Teacher Evaluation Process is to assess the teacher's performance in relation to the North Carolina Professional Teaching Standards and to design a plan for professional growth" (Public Schools of North Carolina, 2015, p. 4). This purpose is further defined by eight goals as follows:

The teacher performance evaluation process will:

1. Serve as a measurement of performance for individual teachers;

2. Serve as a guide for teachers as they reflect upon and improve their effectiveness;

3. Serve as the basis for instructional improvement;

4. Focus on the goals and objectives of schools and districts as they support, monitor, and evaluate their teachers;

5. Guide professional development programs for teachers;

6. Serve as a tool in developing coaching and mentoring programs for teachers;

7. Enhance the implementation of the approved curriculum; and

8. Inform higher education institutions as they develop the content and requirements for teacher training programs. (Public Schools of North Carolina, 2015, p. 5)

To meet these goals, NCTEP incorporates a performance-based evaluation model (Shakman et al., 2012). This comprehensive model includes eight components: (1) training; (2) orientation; (3) self-assessment; (4) preobservations; (5) observations; (6) postconferencing; (7) a summary evaluation conference conducted by the principal, including a summary rating of performance; and (8) professional development plans (Public Schools of North Carolina, 2015). As required by G.S. 115C-333.1(a), all

teachers are evaluated at least annually, and teachers who demonstrate unsatisfactory performance may be placed on mandatory improvement plans (Public Schools of North Carolina, 2015). Annual evaluations are based on one of three evaluation cycles: comprehensive, standard, or abbreviated (Public Schools of North Carolina, 2015).

Given the political climate and public interest in teacher effectiveness, it is likely that the conversation about how to create "a credible system of measuring the quality of teachers' work" (Toch, 2008, p. 1) will continue. Comprehensive, performance-based evaluation systems that include multiple measures of assessment and focus on student growth indicators through VAMs, such as NCTEP, will remain on the rise (Hull, 2013). The task, then, is to capitalize on the processes that lead to fulfillment of the primary goals of teacher evaluation, measuring teacher effectiveness and enhancing professional growth with the hope that "if done well, performance evaluation can be an effective form of teacher professional development" (Taylor & Tyler, 2012, p. 84).

Research Questions

1. What attributes of NCTEP impact teacher perceptions of professional growth?

2. How do the attributes of NCTEP impact teacher perceptions of teacher effectiveness?

3. How does NCTEP impact teacher attitudes toward teaching?

Setting

The setting of this study was a midsized school district located in the foothills of North Carolina. This school district was composed of 26 schools. Among these were 11 elementary schools; four 6-8 middle schools; four K-8 schools; three high schools serving Grades 9-12; an Early College High School and a Middle College High School that served Grades 9-13; and two alternative schools, one that served Grades K-5 and a

second that served Grades 6-12. The Middle College High School and Early College

High School are located on the campus of the local community college; however,

teachers are employed through the county's school system.

The district's Human Resources Department reports that 817 certified teachers

were employed in the school system in the 2015-2016 school year (B. Johnson, personal

communication, April 12, 2016). Johnson stated that teachers are categorized as

elementary (K-8), high school (9-above), or other. Counselors, school psychologists, and

media coordinators are considered certified teachers and are included in the other

category for reporting purposes. The total number of certified teachers is composed of

397 elementary teachers, 198 high school teachers, and 222 teachers in the other category

(B. Johnson, personal communication, April 12, 2016).

At the time of this study, NCTEP had been in use in the district since the 2009-

2010 school year. Training on NCTEP is primarily a responsibility of the Human

Resources Department in this district (T. Johnson, personal communication, January 15,

2016). At the state level, the Educator Effectiveness Division has oversight of NCTEP

(Public Schools of North Carolina, n.d.a). Administrators and teachers have access to

online training modules at the North Carolina Educator Effectiveness System (NCEES)

wikispace (Public Schools of North Carolina, n.d.c). The Educator Effectiveness

Division website is accessible at www.ncpublicschools.org/effectiveness-model/ncees/.

Definition of Terms

Several terms in this study are used synonymously, as this is the manner in which

these terms are discussed in much of the related literature (Darling-Hammond et al.,

1983). Synonymous terms are paired in the list below. Terms are defined as follows.

NCLB. Federal law that mandated testing in reading and math for students in

Grades 3-8 and once in high school. Under NCLB, 100% of students were expected to meet or exceed state proficiency in reading and math by 2014. NCLB aimed to close student achievement gaps by ensuring that all children had opportunities to receive a high-quality education. In its endeavors to achieve this goal, NCLB defined requirements for teachers to be classified as highly qualified and required that student academic growth be definitively measured (NCLB, 2001).

RttT. A federal initiative designed to stimulate educational reform by offering grant incentives to states that were willing to develop and implement systematic plans to improve teaching and learning (American Recovery and Reinvestment Act, 2009). RttT grant competitions focused on setting higher standards, improving teacher effectiveness, using data-driven instruction, and transforming struggling schools. North Carolina was selected to receive RttT grant funding in 2010 (NCLEG, 2010).

NCTEP. The policy adopted by the North Carolina State Board of Education that establishes a comprehensive teacher evaluation with the purpose of assessing "the teacher's performance in relation to the North Carolina Professional Teaching Standards and to design a plan for professional growth" (Public Schools of North Carolina, 2015, p. 4). The NCTEP manual (Public Schools of North Carolina, 2015) stated that

> All of the instruments and processes are designed to encourage professional growth, to be flexible and fair to the persons being evaluated, and to serve as the foundation for the establishment of professional goals and identification of professional development needs. (p. 4)

NCEES. An electronic data collection system that houses all evaluation information on a web-based platform (Lynn et al., 2013; Public Schools of North Carolina, n.d.a). NCTEP is sometimes referred to as NCEES by researchers (Henry &

Guthrie, 2015; Leggett, 2015).

Performance-based teacher evaluation system/comprehensive evaluation system. A system that utilizes multiple measures of teacher performance and employs a variety of evidences to demonstrate teacher knowledge and skills, especially as they relate to student achievement (Hull, 2013; Shakman et al., 2012).

Multiple measures. "Multiple indicators that target a range of components of effective teaching, using such data sources as classroom observations, pre- and post-conference, self-assessments, analysis of classroom artifacts, and professional growth plans" (Shakman et al., 2012, p. 2).

VAMs. Measures of student performance over time as determined through statistical methods that evaluate changes in test scores (Darling-Hammond et al., 2012). EVAAS is North Carolina's state-adopted model for student growth measures relating to school accountability and educator effectiveness (Public Schools of North Carolina, n.d.a). According to SAS (2015), providers of EVAAS, "The goal of value-added reporting is to assess how much students grew and whether that growth is enough to reach a certain goal. . . . These distinctions are important for accountability, evaluation and driving school improvement" (p. 1).

Analysis of student work (ASW). Growth measures used for educators who teach in content areas where standardized tests may not be available or are not appropriate assessments of student performance. Teachers in Arts Education, Healthful Living, World Language, and Advanced Placement and International Baccalaureate teachers collect student work around five objectives that represent their classes and submit work samples into an online platform. Artifacts are evaluated by reviewers who have content area expertise to determine student growth (Public Schools of North

Carolina, n.d.c).

       Status ratings. An overall rating of teachers based on their 3-year rolling average of student growth scores and their summative ratings on the Teacher Evaluation Standards 1-5. The three categories for status are in need of improvement, effective, and highly effective (Public Schools of North Carolina, 2015).

       Student growth. The "amount of academic progress that students make over the course of a grade or class" (Public Schools of North Carolina, n.d.e).

       Professional growth/professional development. Teacher improvement that is fostered by ongoing professional learning that leads to continuous improvement in teaching practices and increased student learning (Darling-Hammond, 2013). Professional development, as defined by the Every Student Succeeds Act (ESSA), refers to the activities that support professional growth (Hirsh, 2015).

       Teacher effectiveness/teacher quality. The impact that the teacher has on students (Darling-Hammond et al., 1983). Medley's (1982) definitions are cited by Darling-Hammond et al. (1983, p. 304) to describe the components of teacher effectiveness, explaining that teacher effectiveness is dependent upon teacher competence (the range of skills, knowledge, and beliefs possessed by the teacher that lead to successful teaching) and teacher performance (what the teacher does in his/her practice in response to specific teaching situations). The expected outcome of effective teaching is student learning (Public Schools of North Carolina, n.d.d). Recently, teacher effectiveness/teacher quality have been closely associated with student growth measures, particulary those determined by VAMs based on standardized test scores (Darling-Hammond et al., 2012; Hull, 2013). The READY document, *Measuring student learning for educator effectiveness: A guide to the use of student growth data in the evaluation of*

North Carolina teachers, explained that "In North Carolina, an effective educator is one who meets the expectations on our Educator Evaluation System" (Public Schools of North Carolina, 2013a, p. 4).

Summary

Regardless of the in-depth study into teacher evaluation, there is no doubt that the issues with rating teacher effectiveness will continue to be problematic (Darling-Hammond, 2014). Although some researchers speculate that performance evaluations can enhance professional growth (Taylor & Tyler, 2012), others postulate that the two purposes of evaluation, measuring teacher effectiveness and stimulating professional growth, are unlikely to coexist in a single evaluation system (Marzano, 2012). Teachers, who are the most affected by the outcomes of evaluation, have only begun to weigh in on how reforms to evaluation processes affect them (Donaldson, 2012).

Chapter 2: Review of Related Literature

Introduction

The purpose of this study was to examine teacher perceptions of the impact that the teacher evaluation process has on professional growth and teacher effectiveness, which have long been identified as the two primary functions of evaluation systems (Danielson & McGreal, 2000). Additionally, the study examined how the evaluation process impacts teacher attitudes toward teaching. Due to political and public pressure, utilizing teacher evaluation as an accountability measure is receiving significant attention (Darling-Hammond, 2014), although advocates of educator evaluation as a means for improving teacher quality remain vocal (Marzano, 2012). Whether it is possible for teacher evaluation to fulfill both summative and formative roles is a controversy that is yet to be resolved (McNergney & Imig, 2003).

Multiple educational research and reform movements parallel the evolution of teacher evaluation systems over the past 50 years (Danielson & McGreal, 2000). Three areas that are of particular significance to educator evaluation include teacher effectiveness research (Darling-Hammond et al., 1983; MET Project, 2013), the standards movement (Danielson & McGreal, 2000), and teacher accountability reforms (Harris et al., 2014). The complex interactions of these elements, combined with policy initiatives such as NCLB and RttT (Darling-Hammond, 2014), have fostered revisions to educator evaluation systems nationwide (Hull, 2013). As part of the new evaluation systems, VAMs, which are determined by student test scores, are often accepted as adequate assessments of teacher effectiveness (Katz, 2016); however, numerous states are facing litigation that alleges unjust personnel decisions have resulted from the use of VAMs as part of the evaluation process (Sawchuk, 2015). In this rapidly changing field, the

ongoing study of teacher evaluation, with particular attention to teacher perspectives, is needed to gauge the capacity to which the system is meeting its primary purposes.

The Historical Context of Teacher Evaluation

The 1800s. From the earliest days of public schools, a focus on the development of effective instructional practices has been considered an important element of teaching. As Marzano et al. (2011) explained,

> The period from the beginning of formal education in the United States up to the mid-1800s saw the dawning of the awareness that pedagogical skills are a necessary component of effective teaching. Although there was little or no formal discussion about the specifics of these skills, the acknowledgment of their importance might be considered the first step in the journey to a comprehensive approach to developing teacher expertise. (p. 13)

Although the particular characteristics of effective teachers and instructional practices were not yet clearly defined, the role of teacher supervision and feedback as a means of improving instruction began to take shape by the mid-1800s (Marzano et al., 2011).

One of the earliest studies attempting to identify desirable characteristics in teachers was conducted by H. E. Kratz in 1896 (McNergney & Imig, 2003). In this study, nearly 2,500 second- through eighth-grade students from Sioux City, Iowa, were asked to identify their best teachers and then describe their characteristics. McNergney and Imig (2003) reported, "Kratz thought that by making desirable characteristics explicit he could establish a benchmark against which all teachers might be judged" (p. 2453). Results of the survey demonstrated that 87% of the students said that helpfulness was the most important characteristic of their best teachers. Rated second at 58%, however, was personal appearance (McNergney & Imig, 2003, p. 2453).

The 1900s-1950s. The principles of scientific management espoused by

Frederick Taylor in the early 1900s, coupled with the measurement theories of Edward

Thorndike, established the basis for a scientific approach to educational processes

(Marzano et al., 2011). According to the principles of scientific management, the

measurement of desired behaviors in factory workers could establish the best method for

completing a task and was an effectual mechanism for increasing productivity. Taylor's

beliefs impacted not only business and industry but public education as is evident in the

works of Ellwood Cubberly and William Wetzel (Marzano et al., 2011).

First published in 1916 and revisited in a third edition in 1929, Ellwood

Cubberly's book *Public School Administration* explained how schools could be viewed

as factories and children seen as the products to be shaped by schools to meet societal

demands (Marzano et al., 2011). Applying principles of measurement to classroom

supervision, Cubberly devised a scale from A to F to rate observed teaching strategies

and to provide a framework for suggested improvements to instruction (Marzano et al.,

2011). William Wetzel, a contemporary of Cubberly, focused less on Cubberly's factory

metaphor but was influenced by his ideas of educational measurement and proposed that

the effectiveness of schools and teachers could be determined through measures of

student learning (Marzano et al., 2011). In addition to focusing on teacher performance

and characteristics, Wetzel (1929) as cited in Marzano et al. (2011),

> recommended three components as the basis for scientific supervision: the use of
>
> aptitude tests to determine the ability level of each child; the establishment of
>
> clear, measureable objectives for each course; and the use of reliable measures of
>
> student learning. (n.p.)

The influence of these ideas is evident in teacher evaluation systems of today (Marzano et

al., 2011; Weisberg et al., 2009).

Interest in the scientific management approach to teacher supervision and effectiveness lost some momentum in the post-World War II era as the progressive education ideals of John Dewey gained ground (Marzano et al., 2011). Dewey's theories sought to promote the ideals of democracy as the basis for educational practices (Dewey, 2014) and advocated for student-centered, differentiated teaching techniques that connected learning to real-world experiences (Marzano et al., 2011). Dewey's philosophies had broad and lasting impact on the field of education; in particular, his theories transformed teacher supervision by fostering the view that the process should include democratic elements and allow for teacher interactions with the administrator (Marzano et al., 2011). Concurrently, approaches to identifying areas of teacher competence took an individualistic turn and focused on the unique characteristics of the teacher (Marzano et al., 2011).

Throughout the 1940s and 1950s, supervisor ratings of teachers were based on presage variables such as teacher traits and personal attributes (Danielson & McGreal, 2000). Specific attributes including voice, appearance, warmth, and enthusiasm were believed to contribute to more effective teaching; however, correlation data reported insufficient evidence to connect these traits to student learning (Danielson & McGreal, 2000). Identifying teacher effectiveness variables confounded researchers, and the preponderance of research indicating that presage variables showed "little or nothing about student learning" (McNergney & Imig, 2003, p. 2453) led some researchers to call "once and for all for an end to research and evaluation aimed at linking teacher characteristics to student learning, arguing it was an idea without merit" (McNergney & Imig, 2003, p. 2453).

As efforts to identify the features of effective teaching evolved (McNergney & Imig, 2003), so did the evaluation process (Marzano et al., 2011). Influenced by Dewey's democratic ideals, instructional leaders received recommendations to recognize individual teacher qualities and to give special attention to their distinctive learning needs as identified during classroom observations. Follow-up conferences which provided opportunities for the exchange of feedback between the teacher and observer were highly encouraged and laid the groundwork for the formative elements of evaluation systems that are intended to promote professional growth (Marzano et al., 2011).

The 1960s-1980s. The belief that teacher effectiveness could be enhanced by fostering teacher growth through observation and feedback encouraged the use of one of the most impactful educational innovations of the postwar era, the clinical supervision model (Marzano et al., 2011). Based on supervisory practices in teaching hospitals, the clinical supervision model, which was defined by Goldhammer (1969), involves a process for teachers and supervisors to participate in reflective conversations to stimulate a teacher's professional growth so that continual improvements to instructional practices and student learning are evidenced (Marzano et al., 2011). Components of the clinical supervision model such as pre and postobservation conferences and classroom observations are standard protocols for the current evaluation systems (Public Schools of North Carolina, n.d.a). However, unlike these recent evaluation models (MET Project, 2013), Goldhammer's supervision model left measures of effective teaching undefined (Marzano et al., 2011).

While the structure of the clinical supervision model became the skeleton for teacher evaluation systems, the holistic approach to observation and the rich dialogues between teacher and observer that Goldhammer envisioned never fully materialized

(Marzano et al., 2011). Rather, the 1960s and 1970s brought new efforts to accurately document what kinds of teacher behaviors could be linked to student achievement (Danielson & McGreal, 2000), and political and public pressures grew to ensure that preservice teachers were well-equipped to begin their careers and that in-service teachers were skilled in delivering effective instruction (Israel, 1978). A new body of work began to grow as teacher effectiveness research, also known as the research on teacher effects, sought to identify effective teaching behaviors as they were linked to student learning (Danielson & McGreal, 2000). Typically in this research, standardized test scores were used as the indicator for the level of student performance; thus, student achievement on standardized tests became a measure of effective teaching (McNergney & Imig, 2003).

The 1980s-present. Madeline Hunter's work in the early 1980s prompted the next phase of transformation in evaluation as her lesson design model provided content for the components of the clinical supervision process: i.e., preconference, observation and postconference (Marzano et al., 2011). Based on a "behavioristic view of basic learning theory" (Danielson & McGreal, 2000, p. 13), Hunter proposed a seven step lesson plan that utilized the learning concepts of motivation, retention, and transfer (Danielson & McGreal, 2000). The seven elements of Hunter's effective lesson included (1) anticipatory set, (2) objective and purposes, (3) input, (4) modeling, (5) checking for understanding, (6) guided practice, and (7) independent practice (Marzano et al., 2011). Although the Hunter model gave teachers guidance in developing student-centered, structured classrooms and was widely accepted by educators, the model was not supported by consistent evidence demonstrating its positive impact on student learning (Danielson & McGreal, 2000). Nonetheless, Hunter's lesson design promoted the generation of evaluation criteria by state and local policymakers that allowed for the

creation of simplistic checklists or rating scales for use as summative measures of teacher performance (Danielson & McGreal, 2000). As reported by Marzano et al. (2011), "Teachers described their lessons in terms of Hunter's model, and supervisors determined the effectiveness of observed lessons in terms of alignment to the model" (p. 20).

Throughout the 1980s and into the 1990s, research into the study of good teaching expanded from Hunter's behaviorist theories to embrace the cognitive learning theory perspective (Danielson & McGreal, 2000). Higher expectations for student performance included more complex learning outcomes, including the attainment of critical thinking and problem-solving skills, and prompted shifts in teaching methodologies and assessments of student learning (Danielson & McGreal, 2000). In relation to teacher effectiveness and evaluation processes, this research continued to support the idea that measuring teacher competency based solely on standardized test scores was problematic and could not capture the full range of professional traits required for effective student learning (Davey, 1991). Paralleling the call for more authentic classroom tests, performance-based assessments were endorsed as more appropriate evaluation tools (Davey, 1991) than checklists (Danielson & McGreal, 2000) or binary rating systems (Weisberg et al., 2009).

An initial step in the modern era of teacher evaluation began with the development of performance-based standards that articulated common goals of the teaching profession by two national organizations, the Interstate New Teacher Assessment and Support Consortium (INTASC, 1992) and the National Board for Professional Teaching Standards (NBPTS). Working collaboratively, these organizations hoped to professionalize education by establishing a shared foundation of the knowledge, skills, and dispositions necessary for effective teaching as well as influencing licensing

and certification practices at the state level (INTASC, 1992) . The work of INTASC and NBPTS in establishing performance-based standards which "describe what teachers should know and be able to do" (INTASC, 1992, p. 7) supported the need for performance-based teacher evaluations (Danielson & McGreal, 2000). As stated by Davey (1991), "Only standards on which there is reasonable professional consensus are fair game for evaluation" (p. 126).

As the body of work in teacher evaluation research increased (Darling-Hammond et al., 1983; Stiggins & Duke, 1988; Wise et al., 1984) so did the pressures for ensuring teacher quality, thus cultivating an atmosphere for the development of educator evaluation systems (Eller & Eller, 2015). Comprehensive evaluation systems, which recognized strengths and weaknesses in evaluatee performance, were promoted as vehicles for meeting both the formative and summative goals of educator evaluation (Gullickson & Howard, 2009). Multiple performance-based measures that documented the actual work of teaching such as portfolios, work samples, and video were included in these comprehensive evaluation systems as they were considered critical to gaining a holistic view of teacher performance (Darling-Hammond, 2013; Davey, 1991). Additionally, technology advances of the mid-1990s gave rise to another innovative component of educator effectiveness measures (Danielson & McGreal, 2000). VAMs, which used statistical methods to analyze student test data in relation to teacher performance, were popularized by William Sanders (Darling-Hammond et al., 2012). Sanders (2000) promoted his valued-added assessments by describing it as a better means for determining whether the gaps in learning gains between diverse groups of students were being effectively closed than "the reporting of simple raw test averages"(p. 332). Furthermore, Sanders held that "a rigorous value-added approach is the fairest, most

objective way to hold districts and schools accountable" (p. 335) and advocated for "value-added estimates of teacher effectiveness [to] be a part of formal teacher evaluation," although cautioning that "they should not be the sole basis upon which teachers are evaluated, because there are too many other duties, dimensions, and responsibilities that cannot be measured by [this] process" (p. 335).

Demands for accountability in teacher quality and evaluation reform span more than 40 years and involve numerous stakeholders, both within and outside the educational community (Darling-Hammond, 2014; Darling-Hammond et al., 1983). In the first decade of the 21st century, federal initiatives including flexibility waivers in NCLB and funding requirements for RttT have continued to fuel the emphasis on measuring teacher effectiveness (Darling-Hammond, 2014) and teacher performance evaluations as a means for improving learning outcomes (Hallinger, Heck, & Murphy, 2014). States have worked to reform their evaluation models to meet these policy demands (Hull, 2013), resulting in new models of teacher evaluation systems that are generally standards-based instruments that employ multiple measures of teacher performance and include a value-added component to measure student learning (Hallinger et al., 2014).

Although there exists an increasing international body of work that demonstrates a direct relationship between teacher quality/effectiveness and student learning, controversy over how best to measure teacher competence and effectiveness remains (Hallinger et al., 2014). While "policymakers increasingly view teacher evaluation as a potentially powerful means of filtering out poor-quality teachers and stimulating instructional improvement among teachers at large" (Hallinger, 2014, p. 6), those embedded in educational research point to the necessity of developing evaluation systems that focus heavily on teacher growth and support teacher professional development

(Darling-Hammond, 2013, 2014; Griffin, 2013; Marzano, 2012). Given these circumstances, it seems likely that the debate will continue as to the probability that an evaluation system can meet the demands necessary to both encourage teacher growth and ensure teacher accountability.

Research and Reform in Teacher Evaluation (from the 1970s to NCLB)

CCSSO Summer Institute, 1978. Literature reveals that the 1970s ushered in an era of unprecedented political activity involving education that resulted in a shift from a focus on civil rights issues in education to a more comprehensive policymaking perspective (Halperin, 1978). These efforts were most evident at the 1978 CCSSO Summer Institute, which was sponsored by CCSSO (Halperin, 1978). Representing 56 entities at this time, these state school officers were responsible for overseeing public education at the state level; and while having different titles and different means of appointment to their positions, the members shared the common goal of improving education (Campbell, 1978). Choosing to address political issues in a variety of educational arenas, the 1978 Summer Institute demonstrated the desire for CCSSO to "do more by working together to become a political force in this period of transition [and] to chart the course of change" (Hechinger, 1978, p. 134).

Presenters at the 1978 CCSSO Summer Institute spoke at length about the quality of public education, and a consensus existed that national collaboration on these issues would result in better solutions as well as more credibility for teaching as a profession (Israel, 1978). Speaking on the subject of establishing national education priorities, Halperin (1978) directed attention to the growing level of activism by special interest groups and to political initiatives aimed at creating a national dialogue for the establishment of a common strategy to strengthen policymaking in education. Halperin

explained that "this policymaking in education tends to be far more vigorous/activist under the Democrats" (p. 17) but credited a much less partisan atmosphere with cultivating a climate conducive for discussions of "pervasive and significant educational issues" (p. 5). Given these unified efforts to address educational issues, Halperin suggested five strategies that he believed would legitimize national policymaking in the field of education: (1) development of a cabinet-level Department of Education which would become law in 1979; (2) a unified state front, as the CCSSO was currently exhibiting; (3) formation of coalitions with noneducational groups; (4) making hard choices and choosing areas of focus for policymaking; and (5) self-regulation and preemption by educators.

Halperin's last recommendation, that educators develop procedures for tackling their own problems through self-regulation and preemptive actions, was repeated by other speakers throughout the conference, notably in sessions on competency testing and teacher education and certification (Brain, 1978; Brickell, 1978). Brickell (1978) expanded the discussion of political and public involvement in educational affairs by identifying a long list of reasons why demands for competency testing of preservice teachers were valid. Brickell asserted, "*The public was past ready. And it had more than enough reasons*" (p. 69); and "*Legislators were quite sensitive. And they had excellent reasons*" (p. 70). Some of the reasons for the public's demands were identified by Brickell as

- Standardized test scores, including SAT and NAEP, were in decline.
- Employers reported inadequately prepared employees.
- Unemployment rates were rising.

- Students demanded higher standards.

- There was no closing of the achievement gap between advantaged and disadvantaged students.

- School spending was increasing without achieving adequate results,

- "A creeping suspicion that teachers are not as good as they used to be" was prevalent. (Brickell, 1978, p. 90)

Legislative backing for competency testing stemmed from public opinion concerns, with legislators recognizing that "money for education makes up most of the state budget" (Brickell, 1978, p. 70) and that their stand on this issue could affect their political careers (Brickell, 1978). Brickell (1978) commented that legislators recognized the benefits of competency testing for preservice teachers as an accountability measure and delineated other organizations that were for, against, or neutral in the debates over competency testing. Pointing to the role of CCSSO, Brickell stated that their stand, which recommended that a level of competence be defined and measured, put them in the role of "translating, mediating" (p. 74) between the profession and the citizens' demands for accountability (Brickell, 1978). Brickell concluded his statements by suggesting that a related idea may be brewing to ensure teacher accountability, referencing recommendations of the vice-chancellor of the New York State Board of Regents on July 28, 1978, "that minimum competency standards be developed for teachers to insure that they were doing a proper job in the classroom" (p.77).

Brain (1978), Dean of the College of Education at Washington State University, addressed political issues in teacher evaluation and certification during the 1978 CCSSO Summer Institute. Brain's statements echoed those of Halperin (1978) and Brickell

(1978) in reinforcing the thought that educators needed to take a critical look at the state of their field. Brain stated that negativity toward the quality of teaching "has led legislators and congressional people to focus their attention on preparation programs at colleges and universities" (p. 119). Brain suggested that reform in teacher preparation programs was necessary and could be done more effectively due to the past decade's research into preservice programs. Some lessons Brain said had been learned through this research included that

- Teacher development takes time.

- Student teaching should include "frequent, varied and criticized practice: observation, immediate feedback and practice again" (Brain, 1978, p. 121).

- Collaboration between the state, LEAs, and universities must occur.

- Teachers need skills training in humanistic and behavioral sciences.

- Competency standards should be set through certification requirements and accreditation standards.

- "principle of individual differences applies to teachers and to teacher training as well as to children" (Brain, 1978. p. 122).

- Preservice and in-service teachers have much in common and can benefit from ongoing professional development.

- "Excellent teacher training requires budget support" (Brain, 1978, p. 123).

Brain (1978) called for more involvement by the CCSSO in preservice issues, stating, "Chief state school officers typically have not seen teacher education and certification as a political concern of state departments of education" (p. 124). However, Brain believed that political forces should work together to make "the next decade . . .

one of the most successful in the history of teacher education and administrator preparation," and that changes in programs will mean " the political clout of educators can be enhanced because the public will have confidence in their processes and their products" (p. 125).

Hechinger (1978) , president of the New York Times Foundation, closed the Summer Institute with a challenge to education leaders to support educational reforms. Hechinger said,

> There is dissatisfaction with the quality and the effectiveness of our educational system and dissatisfaction over the relationship between the amount of money that is spent and the return on that expenditure. The "back to basics" movement is real. It springs from a variety of motivations. Some support it only because they think a reduction to the basics is cheaper; but others believe that an emphasis on the basics is educationally more productive. The conclusion differs little from my observations about educational politics and the courts; it is up to the schools to educate, to take the lead and to define the basics. (p. 131)

Hechinger's call depicts the pressures faced by educators to prove to the public that they were providing a high-quality education to all students.

Darling-Hammond et al. (1983). As the education field attempted to respond to the increasing demands for teacher accountability, teacher evaluation systems were considered as a means to meet this goal (Darling-Hammond et al., 1983; Stacey, Holdzkom, & Kuligowski, 1989). While the publication of *A Nation at Risk* by the National Commission on Excellence in Education in 1983 provided a summary of the need for education reform, numerous stakeholders were already involved in the research and development of evaluation systems (Stacey et al., 1989). Published in the fall of

1983, *Teacher Evaluation in the Organizational Context: A Review of the Literature* (Darling-Hammond et al., 1983) examined "research on teaching, organizational behavior, and policy implementation" (p. 285) to determine the theories that underlie evaluation models. Sponsored by the National Institute of Education, in collaboration with the RAND Corporation, the group's research explored teaching effectiveness and the organizational context of evaluation models as they existed within the demands of policy constructs (Darling-Hammond et al., 1983).

In relation to the policy constructs of teacher evaluation, this literature review reported policy changes that set higher standards in the areas of certification, evaluation and tenure of pre and in-service teachers, although the report found few school districts that were using evaluation results for personnel decisions in that era (Darling-Hammond et al., 1983). Nonetheless, the consideration of due process procedures impacted evaluation models. Beckham (1981), as cited by Darling-Hammond et al. (1983), recommended that evaluation policies include (1) a predetermined standard of teacher knowledge, competencies, and skills; (2) an evaluation system capable of detecting and preventing teacher incompetencies; and (3) a system for informing teachers of the required standards and according them an opportunity to correct teaching deficiencies (p. 287).

Application of these criteria to evaluation systems was predicted to create difficulties in reconciling the summative and formative purposes of evaluation. Darling-Hammond et al. (1983) stated, "The most obvious problem is that developing a predetermined standard of teacher knowledge, competencies, or skills poses nontrivial controversies about the content and specificity of the standards" (p. 287). Additionally, Darling-Hammond et al. (1983) noted that the use of standardized tests or performance

assessments for job status decisions must be valid, demonstrating a direct relationship to effective teaching. In essence, they argued that "the context-free generalization necessary for implementing a uniform evaluation system may counteract the context-specific processes needed to effect change in individual or organizational behaviors" (Darling-Hammond et al., 1983, p. 288).

A review of effective teaching research, as discussed by Darling-Hammond et al. (1983), attempted to link teacher behaviors to increased student learning; however, this task was complicated by conflicting conceptualizations of teaching work. Whereas some concepts of the teaching work align to a uniform evaluation approach and "effective practices can be determined and specified in concrete ways" (Darling-Hammond et al., 1983, p. 291), other conceptions of teaching require "a repertoire of specialized techniques [and] the exercise of judgment about when those techniques should be applied" (Darling-Hammond et al., 1983, p. 291). Based on inconsistencies in the teaching effects research and varying conceptions of the teaching work, the authors postulated that clear definition of educational goals would be a significant step in the development and design of teacher evaluation systems (Darling-Hammond et al., 1983).

Two divergent teacher evaluation systems emerged as a result of the perceptions of teaching (Darling-Hammond et al., 1983). One model, embraced by many states, was a competency- or performance-based system that assumed the "validity, stability, and generalizability of effective teaching behaviors" (Darling-Hammond et al., 1983, p. 299). The alternative model employed clinical supervision methodologies, self-assessments, and interaction between the observer and teacher to determine if teacher performance reflected an appropriate response to the specific learning situation. Researchers proposed that the beliefs about how schools should operate were intricately connected to the type

of teacher evaluation system chosen for implementation (Darling-Hammond et al., 1983).

The basic purposes of teacher evaluation, which are improvement and accountability, were described in the literature review and were viewed at both the individual and organizational levels (Darling-Hammond et al., 1983). Table 1 illustrates the purposes of evaluation at the individual and organizational levels (Darling-Hammond et al., 1983, p. 302).

Table 1

Four Basic Purposes of Teacher Evaluation

Purpose/Level	Individual	Organizational
Improvement (formative information)	Individual Staff Development	School Improvement
Accountability (summative information)	Individual personnel (job status decisions)	School status (e.g., certification) decisions

Darling-Hammond et al. (1983) indicated that "Many teacher evaluation systems are nominally intended to accomplish all four purposes, but different processes and methods are better suited to one or another of these objectives" (p. 302). For instance, staff development is acknowledged as an important component of evaluation systems in that it is necessary for individual and organizational improvement. However, the authors noted that when the purpose for staff development shifted from improvement to accountability, the usefulness of the activity may be impacted (Darling-Hammond et al., 1983).

Contrasting the purposes of evaluation, Darling-Hammond et al. (1983) explained,

In general, teacher evaluation processes most suited to accountability purposes

must be capable of yielding fairly objective, standardized, and externally

defensible information about teacher performance. Evaluation processes useful

for improvement objectives must yield rich, descriptive information that

illuminates sources of difficulty as well as viable courses for change. Teacher

evaluation methods designed to inform organizational decisions must consider the

context in which individual performance occurs to ensure appropriateness and

sufficiency of data. (p. 303)

Adding further commentary, Darling-Hammond et al. (1983) continued by saying,

Although these purposes and the approaches most compatible with them are not

necessarily mutually exclusive, an emphasis on one may tend to limit the pursuit

of another. Similarly, while multiple methods for evaluating teachers can be used

– and many argue, should be used – it is important to consider what purposes are

best served by each if teacher evaluation goals and processes are to be consonant.

(p. 303)

A number of the general evaluation processes and methods identified in this early

literature review (Darling-Hammond et al., 1983) are commonly recognizable in recent

evaluation trends (Hull, 2013). These are (1) teacher interviews/conferences, (2)

competency testing, (3) indirect measures such as presage variables and professional

experience, (4) classroom observation, (5) student ratings, (6) peer review, (7) student

achievement, and (8) self-evaluation. Darling-Hammond et al. (1983) ascertained that

the effectiveness of these methods is dependent on the goals of the measurement system

and that none should be considered as adequate to "capture enough information about

teaching attributes to completely satisfy any of the purposes for teacher evaluation" (p.

308).

By the early 1980s, literature demonstrates that numerous models of evaluation systems were in use (Darling-Hammond et al., 1983). While unique in design, these evaluation systems were similar in the fact that the manner in which the evaluation methods were configured within the systems was responsive to three criteria:

- What teacher attributes (e.g., professional training, teaching competencies, etc.) are believed to be important for effective teaching;

- Which aspect of the instructional process the district hopes to affect (e.g., assurance of teacher quality; improved teaching techniques; learning outcome, etc.); and

- What will be the criteria for evaluating success (e.g., demonstration by the teacher of desired behaviors or competencies, teacher or student test scores, etc.). Darling-Hammond et al. (1983, p. 309)

Organizational goals and values were revealed as impacting the application of the evaluation system, and the beliefs of the individual participants in the process were shown to enhance or hinder implementation of the system. Darling-Hammond et al. (1983) summarized,

> Effective change requires a process of mutual adaptation in which change agents at all levels can shape policies to meet their needs – one in which both the participants and the policy are transformed by the convergence of internal and external reference points. (p. 317)

In the culmination of *Teacher Evaluation in the Organizational Context: A Review of the Literature,* four minimal conditions were presented for a successful

educator evaluation system:

- All actors in the system have a shared understanding of the criteria and processes for teacher evaluation;

- All actors understand how these criteria and processes relate to the dominant symbols of the organization, that is, there is a shared sense that they capture the most important aspects of teaching, that the evaluation system is consonant with educational goals and conceptions of teaching work;

- Teachers perceive that the evaluation procedure enables and motivates them to improve their performance; and principals perceive that the procedure enables them to provide instructional leadership;

- All actors in the system perceive that the evaluation procedure allows them to strike a balance 'between adaptation and adaptability, between stability to handle present demands and flexibility to handle unanticipated demands' (Weick, 1982, p. 674); that is, that the procedure achieves a balance between control and autonomy for the various actors in the system. (Darling-Hammond et al., 1983, p. 320)

As referenced above, these authors noted that their review of the literature did not demonstrate that the aforementioned conditions for an effective evaluation system existed and cautioned their readers that more work in the field of evaluation was needed before evaluation could contribute to teacher improvement (Darling-Hammond et al., 1983).

Wise et al. (1984). Shortly after *Teacher Evaluation in the Organizational Context: A Review of the Literature* was published, a second report was presented by the RAND group in response to growing concerns about teacher quality (Wise et al., 1984).

Also, commissioned by the National Institute of Education, this report was published in

1984 and focused on the actual practices and procedures employed in evaluations,

including the instruments utilized and the context of the organization in which

evaluations were conducted as well as how evaluation results were used by the

organization (Wise et al., 1984). Noting concerns that merit pay, career ladder, and

master teacher policies should be tied to reliable, effective evaluation systems, Wise et al.

(1984) emphasized the importance of effective evaluation practices, stating,

> School district administrators must understand the educational and organizational
>
> implications of the teacher evaluation system that they adopt, because that system
>
> can define the nature of teaching and education in their schools. In particular, the
>
> system can either reinforce the idea of teaching as a profession, or it can further
>
> deprofessionalize teaching, making it less able to attract or retain talented
>
> teachers. (p. v)

This RAND study utilized a literature review and an examination of four districts

with highly developed evaluation systems as its basis (Wise et al., 1984). The evaluation

systems selected for investigation were determined successful by the researchers if they

were implemented as planned, had stakeholders who demonstrated a common

understanding of processes, and informed teacher improvement and personnel decisions

through results that were valid and reliable. Major issues related to principals, teachers,

and systems were consistently identified across the evaluation systems studied.

Participants commented on the principals' lack of competence and resolve to conduct

accurate evaluations and noted, "other problems included teacher resistance or apathy,

the lack of uniformity and consistency of evaluation within a school system, inadequate

training for evaluators, and shortcomings in the evaluation of secondary school staff and

specialists" (Wise et al., p. vi). On a positive note, respondents indicated that the

evaluation process improved teacher/administrator communication and focused teachers'

attention on instructional goals and teaching practices. Generally, teachers felt a sense of

professionalism associated with their participation in the evaluation process (Wise et al.,

1984).

By examining the organizational characteristics of the districts in this study in

relation to the success of the adopted teacher evaluation systems, researchers were able to

generate a set of five conclusions and 12 recommendations to serve as strategies for

effective implementation of teacher evaluation models (Wise et al., 1984). Conclusions

and recommendations are reported in Table 2.

Table 2

Conclusions and Recommendations for Effective Implementation of Teacher Evaluation Models (Wise et al., 1984. pp. xi-xiii)

Conclusion One:

> *To succeed, a teacher evaluation system must suit the educational goals, management style, conception of teaching, and community values of the school district.*

Recommendations:
1. The school district should examine its educational goals, management style, conception of teaching, and community values and adopt a teacher system compatible with them. It should not adopt an evaluation system simply because that system works in another district.
2. States should not impose highly prescriptive teacher evaluation requirements.

Conclusion Two:
> *Top level commitment to and resources for evaluation outweigh checklists and procedures.*

Recommendations:
3. The school district should give sufficient time, unencumbered by competing administrative demands, for evaluation. This may mean assigning staff other than the school principal to some evaluation functions.
4. The school district should regularly assess the quality of evaluation, including individual and collective evaluator competence. The assessments should provide feedback to individual evaluators and input into the continuing evaluator training process.
5. The school district should train evaluators in observation and evaluation techniques, including reporting, diagnosis, and clinical supervision skills, when it adopts a new teacher evaluation process.

Conclusion Three:
> *The school district should decide the main purpose of its teacher evaluation system and then match the process to the purpose.*

Recommendations:
6. The school district should examine its existing teacher evaluation system to see which, if any, purpose it serves well. If the district changes the purpose, it should change the process.
7. The school district should decide whether it can afford more than one teacher evaluation process or whether it must choose a single process to fit its main purpose.

Conclusion Four:

> *To sustain resource commitments and political support, teacher evaluation must be seen to have utility. Utility depends on the efficient use of resources to achieve reliability, validity, and cost-effectiveness.*

Recommendations:

8. The school district must allocate resources commensurate with the number of teachers to be evaluated and the importance and visibility of evaluation outcomes.
9. The school district should target resources so as to achieve real benefits.

Conclusion Five:

> *Teacher involvement and responsibility improve the quality of teacher evaluation.*

Recommendations:

10. The school district should involve expert teachers in the supervision and assistance of their peers, particularly beginning teachers and those in need of special assistance.
11. The school district should involve teacher organization in the design and oversight of teacher evaluation to ensure its legitimacy, fairness, and effectiveness.
12. The school district should hold teachers accountable to standards of practice that compel them to make appropriate instructional decisions on behalf of their students.

Stiggins and Duke (1988). Interest in improving teacher evaluation systems continued to grow throughout the 1980s (Stiggins & Duke, 1988). Stiggins and Duke (1988), stipulating that "teacher evaluation is a key to school improvement" (p. xi) but that "evaluations have failed to improve teaching" (p. xi), conducted four studies over a 3-year span that were designed to identify how teacher evaluation could be reformed "to promote teacher growth and development" (p. xi). Their work began with a literature review of teacher evaluation followed by an examination of four districts' evaluation systems and continued with interviews of teachers who reported participating in evaluation processes that supported professional growth (Stiggins & Duke, 1988). Finally, the researchers utilized the Teacher Evaluation Profile (TEP) Questionnaire to

survey 470 teachers from across five school districts to determine teacher perceptions of their experiences with the evaluation process (Stiggins & Duke, 1988).

The guiding premise of Stiggins and Duke's (1988) research was that the best outcome of an evaluation system was when teacher evaluation fostered professional growth. Citing Stiggins' (1986) earlier work, Stiggins and Duke (1988) commented,

Accountability systems strive to affect school quality by protecting students from incompetent teachers. However, because nearly all teachers are at least minimally competent, the accountability system directly affects only a very few teachers who are not competent. Thus, if our goal is to improve general school quality – and we use only those strategies that affect a few teachers – overall school improvement is likely to be a very slow process.

Growth-oriented systems, on the other hand, have the *potential* of affecting all teachers-not just those few who are having problems. There is no question that all teachers can improve some dimension(s) of their performance. (Stiggins, 1986, pp. 53-54)

Stiggins and Duke (1988) noted that the personal motivation necessary for teacher improvement is impacted by the purposes of the evaluation system, whereas an accountability model compels teachers to change through threat of employment status, a model that promotes professional development values intrinsic teacher motivation.

In Stiggins and Duke's (1988) review of the literature, common themes emerged as to the state of evaluation. Work by Lewis (1982) revealed that both administrators and teachers had concerns about the processes of teacher evaluation systems and saw only weak links between evaluation and its impact on instructional improvements, unclear performance standards and the interpersonal dimensions of evaluation as obstacles to the

process. Other researchers in the Stiggins and Duke's (1988) literature review offered

suggestions for revisions of evaluation systems. These researchers encouraged the use of

multiple measures of teacher performance such as student and peer reviews and ASW

products, along with additional means of data collection (i.e., charting/recording

classroom interactions) in the evaluation process (O'Hanlon & Mortenson, 1977; Stiggins

& Duke, 1988). Components of exemplary evaluation programs were also described in

the literature review (Stiggins & Duke, 1988); and McLaughlin (1982) reported the most

common features of these programs as observations of classrooms, pre and

postobservation conferences, and documented action plans with follow-up by the

principal who acted as the primary evaluator. However, Stiggins and Duke summarized

that even in exemplary programs, the evaluation process is not integrated into an overall

staff development plan.

Following their literature review, Stiggins and Duke (1988) devised three studies

to determine the factors impacting the potential for an evaluation system to promote

teacher growth. Their first study documented specific elements of four teacher

evaluation systems using a case study methodology. Conclusions generated from teacher

and supervisor perceptions identified numerous changes that were needed for more

effective evaluations:

- Methods should match purposes of evaluations.

- Teachers should be actively involved with all phases of an evaluation system,
  from development to the operational stages.

- Thorough training on the system should occur for all participants.

- Multiple measures of data should be part of the evaluation.

- Performance criteria should be clear, appropriate and connected to anticipated student outcomes.

- Results of observations should be used in a meaningful way. (Stiggins & Duke, 1988)

Stiggins and Duke (1988) also acknowledged an overarching assumption that is necessary for implementation of a growth-based evaluation system, stating, "teachers and supervisors alike function best within an environment characterized by mutual support, and by respect and concern for personal growth and for the wellbeing of staff members and students" (p. 24).

A second case study by Stiggins and Duke (1988) examined the experiences of teachers and supervisors who had been involved in an evaluation that they perceived as leading to professional growth. Using an analysis of verbatim interview responses, the researchers were able to describe specific attributes of these evaluations that could contribute to an evaluation experience that promoted teacher improvement (Stiggins & Duke, 1988), although they noted that "because of the complexity of human relations in the evaluation context, we cannot predict with certainty which evaluations will or will not be helpful to teachers" (Stiggins & Duke, 1988, p. 79). They surmised that "the probability a teacher will benefit from an evaluation increases as the number of these attributes present in the evaluation increases" (Stiggins & Duke, 1988, p. 79).

Five key factors that contributed to the quality and impact of an evaluation experience emerged from Stiggins and Duke's (1988) studies. These factors, designated as important attributes, with their descriptors are

1. Important Attributes of the Teacher

    a.   Instructional Competence

    b.   Personal Expectations

    c.   Openness to Suggestion

    d.   Orientation to Change

    e.   Subject Knowledge

    f.   Experience

2. Important Attributes of the Person Who Observes and Evaluates

    a.   Credibility

    b.   Persuasiveness

    c.   Patience

    d.   Trust

    e.   Track Record

    f.   Modeling

3. Important Attributes of the Procedures Used to Gather Data on Teacher

Performance

    a.   Performance Criteria and Standards

    b.   Data Collection Procedures

4. Important Attributes of the Feedback

    a.   Appropriate communication of information delivered in respectful manner

5. Important Attributes of the Evaluation Context

    a.   Labor Relations

    b.   Time Spent on Evaluation

    c.   Resources Available for Growth (Stiggins & Duke, 1988).

The success of these attributes in producing evaluation results that stimulated

professional improvement encouraged Stiggins and Duke to determine,

> The key to success, based on our limited number of cases, appears to be careful attention to *all* of the active ingredients in the evaluation process: teachers and supervisors must be willing to contribute to that success; procedures for gathering performance data must be carefully planned and carried out; feedback must be delivered in a thoughtful manner; and the context must be one in which teacher growth is valued. (p. 92).

To validate these conclusions, a third study, which utilized a questionnaire to gather information from a large sampling of teachers about their most recent evaluations, was developed by the researchers (Stiggins & Duke, 1988). TEP asked teachers to rate and describe the quality and impact of their evaluation experience. The five key factors identified in their previous study provided the basis for the creation of TEP, and results were utilized to generalize the findings from their case studies to other teachers and their experiences. Stiggins and Duke (1988) reported that the data do indicate a strong "relationship between the attributes of a teacher evaluation and its perceived impact on teachers" (p. 117) as well as highlighting some of the most crucial attributes of a growth-based evaluation system. In conclusion, the authors indicated that additional studies that employ TEP should produce useful information for evaluation reform (Stiggins & Duke, 1988).

Teacher evaluation policies for beginning teachers. Beginning teacher evaluation policies were described in work by Sclan and Darling-Hammond (1992). Reflecting on the significance of beginning teacher evaluations, Sclan and Darling-Hammond proposed that beginning teachers should be provided with

> a collaborative growth-oriented environment . . . intended to perpetuate an active

style of teaching in which teachers seek to discover and use an ever-expanding set

of strategies to meet student needs, rather than relying on a limited set of routines

regardless of their effectiveness with particular students. (p. 13)

Confirming the close connection between policy making and evaluation reform, the

authors explained that state involvement in programs for beginning teacher supervision

and evaluation increased rapidly between 1988 and 1992, noting that 45 states enacted

beginning teacher requirements by 1992. However, in reaction to funding cutbacks, they

also indicated that several states had recently either slowed or suspended these programs

(Sclan & Darling-Hammond, 1992).

Teacher evaluation and student achievement measures. As discussions of

how to develop and implement effective teacher evaluation systems continued, the use of

multiple data sources as evaluative measures was considered by researchers (Darling-

Hammond et al., 1983; Stiggins & Duke, 1988). In particular, deliberations on the use of

student achievement data as a teacher accountability measure raised questions about the

appropriateness of standardized test scores in teacher evaluation (Sanders & Horn, 1998;

Stiggins & Duke, 1988). In Haertel (1986), student achievement test scores were viewed

as plausible indicators of teacher effectiveness, especially in identifying poor teacher

performance. However, the author cautioned that if tests were used incorrectly, biases in

favor of teachers with higher performing students could occur (Haertel, 1986).

Sanders (2000), a statistician for the SAS Institute, was one of the most vocal

spokespersons for the inclusion of VAMs in teacher evaluation systems (Easton, 2012).

Sanders's model, also referred to as a growth-based model, used a comparison of student

current test scores to their previous scores in order to make predictions about future

achievement levels. Student academic growth rates could then be considered as

indicators of teacher and school effectiveness (Easton, 2012).

Speaking at the Consortium for Research on Educational Assessment and Teaching Effectiveness (CREATE) National Evaluation Institute on July 21, 2000, Sanders presented his rationale for the use of value-added assessments and standardized test data as measures of teacher and school effectiveness. Sanders (2000) stated that "In the last years of the twentieth century, education was called to account for the failures of large segments of the population of American students to achieve the minimal academic expectations of society as a whole" (p. 329) and expounded on how past educational philosophies and strategies had failed to eliminate inequities in the academic achievement. Sanders further explained that while proficiency measures varied across student demographics, "rates of academic progress can be estimated nearly, if not entirely, free of socio-economic and ethnic confoundings" (Sanders & Horn, 1998, p. 331). Sanders reasoned that his value-added model monitored the speed at which student growth occurred and, thus, would be a more appropriate measure of the effectiveness of instruction. Sanders said,

> An accountability system that is based upon the rate of academic progress of populations of students is one that will hold people accountable for things over which they have control, rather than for things over which they do not. For instance, teachers in the fall have no control over the achievement level of their incoming students. However, teachers do have primary control over the achievement rate of academic progress of their students. (p. 331)

The Tennessee State Board of Education was among the first public school entities to embrace Sanders' (2000) methods for determining teacher and school effectiveness by implementing the Tennessee Value-Added Assessment System, or

TVAAS, as part of the state's teacher evaluation system (Easton, 2012). NCLB (2001) fostered additional support for Sanders' theories when the U.S. Department of Education approved the growth model as a means to determine that schools were achieving Adequate Yearly Progress (AYP) per mandates of the 2002 law (Easton, 2012).

The Impact of Federal Policymaking on Teacher Evaluation in the 21st Century

Stronge and Tucker (1999) addressed the complex political aspects involved in the implementation of new evaluation systems. Acting as participant-observers, Stronge and Tucker studied one school district's efforts to revise its evaluation system so evaluation was linked to personnel performance based on school and district goals. Over 3 years of observations, Stronge and Tucker identified crucial elements of the political climate that contributed to effective organizational change: (1) systemic compatibility between the evaluation system and other parts of the educational system such as the improvement process and professional development; (2) effective, ongoing communication between all stakeholders, particularly teachers; (3) an organizational commitment of resources including time; and (4) collaboration between teachers and administrators.

Findings from Stronge and Tucker's (1999) study stressed the importance of involving teachers in the development and implementation of the evaluation system so they experienced a sense ownership, trust, and respect in the process. Stronge and Tucker speculated that by meeting the needs of the individuals in the system as well as the needs of the institution, an evaluation system could enhance teacher and school effectiveness. However, the researchers also noted that the true organizational change, in which teacher evaluation becomes a tool for growth and improvement, could not be hurried (Stronge & Tucker, 1999).

Other studies of the late 1990s examined the impact of political demands on evaluation (Bridges & Groves, 1999; Lofton, Hill, & Claudet, 1997). Based on their in-depth study of the initial implementation of a performance-based evaluation system, Lofton et al. (1997) questioned whether state-mandated evaluation systems could promote school improvement and improve professional development. They concluded that "In the complex world of teaching and learning, however, there are *no quick fixes. Substantive changes that focus on the teacher as decision maker and reflective practitioner require time and energy*" (Lofton et al., 1997, p. 158). Furthermore, they suggested that "growth and improvement were more influenced by factors influencing learning in general than by traditional measurement and evaluation issues" (Lofton et al., 1997, p. 158.). While still supporting the reform of evaluation systems, Lofton et al. noted that there is a "need to *go beyond the assessment system itself to examine issues related to using that system to improve teaching and learning and to effect changes in the everyday life of the school*" (p. 158).

Bridges and Groves (1999) took the position that "politics plays a major, often detrimental, role in personnel evaluation" (p. 321). Through their research into the macro and micropolitics of evaluation, these authors determined that "viewing personnel evaluation simply in rational, technical terms conceals more than it illuminates" (Bridges & Groves, 1999, p. 336). They proposed that the conflicting interests of stakeholders can complicate the utility of teacher evaluation as a tool for improving student performance (Bridges & Groves, 1999).

Nonetheless, the emphasis on teacher evaluation as a tool for school improvement has intensified in the first decade of the 21st century and, according to seasoned researchers, has often left little time for reflection on the research base of previous

decades (Darling-Hammond, 2013). Although the Bush administration's NCLB (2001) did not implicitly require changes to educator evaluation systems, its focus on ensuring that teachers were highly qualified fueled ongoing debates about how to define teacher quality (Lyttle, 2011). Additional NCLB mandates required that student academic growth be definitively measured (Easton, 2012; NCLB, 2001) and that stiff sanctions be imposed on schools that failed to demonstrate the academic progress of its students (Toch, 2008). These mandates, in addition to the requirements of NCLB flexibility waivers, impacted teacher evaluation (Darling-Hammond, 2014) and prompted the adoption of VAMs into state evaluation systems (Easton, 2012).

Obama's RttT initiative, funded through the American Recovery and Reinvestment Act of 2009, had even more bearing on teacher evaluation than NCLB (Pathe & Choe, 2013). This federally-funded competition explicitly called for reforms in teacher evaluation systems as a means of improving education and established five criteria for state grant proposals involving evaluation (Databases on State Teacher and Principal Evaluation Policies, n.d.). The criteria required states to

    (1) Establish clear approaches to measuring student achievement growth for individual students.

    (2) Design and implement rigorous, transparent, and fair evaluations for the teacher.

    (3) Differentiate effectiveness using multiple rating categories that take student achievement growth into account as a significant factor and are designed with teacher involvement.

    (4) Conduct annual evaluations that include timely and constructive feedback and provide teachers with data on student achievement growth for their students,

classes, and schools.

(5) Use evaluations to inform decisions about staff development, compensation,

promotion, tenure, certification, and removal of ineffective teachers.

(National Center for Educational and Regional Assistance, 2014, p. 2)

In search of a share of the approximately $4.35 billion designated for school

improvement by RttT (Pathe & Choe, 2013), a majority of states developed and

implemented new teacher evaluation systems between 2009 and 2013 (Hull, 2013), yet

little is certain about the progress that states have made toward policy changes (National

Center for Educational Evaluation and Regional Assistance, 2014) or how utilizing

student achievement as a measure of teacher effectiveness impacts teacher evaluations

(Hull, 2013).

Teacher Evaluation in North Carolina

The development of teacher evaluation in North Carolina mirrors evaluation

reform in the United States (Stacey et al., 1989). As noted by Stacey et al. (1989), even

before the publication of *A Nation At Risk* in 1983, politicians and other stakeholders

were considering education reforms that would allow "America to maintain its

preeminent position in the global community" (p. 79). Emphasizing North Carolina's

advances, Stacey et al. remarked,

> Since 1978, for example, the General Assembly of the State of North Carolina has
>
> been systematically addressing the needs – both in policy and fiscal arena – for
>
> education improvements with respect to facilities; texts and text selection
>
> procedures; and certification, salary and benefits, and evaluation of educators. (p.
>
> 79)

Development of North Carolina teacher evaluation systems. The earliest

efforts to develop an instrument for teacher evaluation faced hurdles, however, as research about the skills necessary for effective teaching was unavailable during that time (Stacey et al., 1989).   In an attempt to overcome this deficit, teachers, principals, and superintendents were asked to identify teacher responsibilities. A factor analysis of these practices led to the creation of a consensus-based teacher evaluation instrument (Stacey et al., 1989).

Several problems existed with this first evaluation instrument (Stacey et al., 1989). As reported by Stacey et al. (1989), studies by Inman (1982a, 1982b) revealed that individual principal ratings of teachers were similar across multiple dimensions, and interrater reliability was questionable as overall ratings on the quality of performance varied with different observers (Stacey et al., 1989). Additionally, Stacey et al. noted that the instrument primarily served a summative function, rather than encouraging instructional improvements, and serious issues were present even for the summative functions, including

> (1) The criteria or standards themselves were not clearly understood by all
>
>   observers;
>
> (2) Procedures for data collection and analysis varied from school to school; and
>
> (3) Use of multiple observers was not required. (p. 80)

These concerns led to the development of a second evaluation instrument known as the Teacher Performance Appraisal Instrument (TPAI; Stacey et al., 1989). Drawing from the increasing body of research into teacher and school effectiveness, a contracted group of university researchers, in conjunction with teachers and administrators, reviewed relevant literature to identify 28 teaching skills. These skills were categorized in five teaching areas which made up the core of the evaluation instrument: (1)

management of instructional time, (2) management of student behavior, (3) instructional presentation, (4) instructional monitoring, and (5) instructional feedback (Stacey et al., 1989). Three additional areas were included to represent the values and norms of the organization, although these were not derived from the literature review. These were (6) facilitating instruction, (7) interacting within the educational environment, and (8) noninstructional duties. These eight functions were further defined by 38 practices (Stacey et al., 1989). Ratings for the evaluation were given at the function level, rather than the practice level (Holdzkom, 1987). The full TPAI with eight functions and 38 practices is presented in Appendix A.

Procedures for use of TPAI were established based on two purposes: (1) providing ongoing feedback and promoting professional development for teachers and (2) producing defensible summative ratings for personnel decisions (Stacey et al., 1989). An evaluation cycle was established that included "1) multiple classroom observations, 2) observations that focused on a variety of teaching behaviors of high and low inference, and 3) documentation that would result in periodic narrative formative evaluation reports and a year-end quantified, summative evaluation report" (Stacey et al., 1989, p. 81; Holdzkom, 1987). In order to foster a climate conducive to instructional improvement, evaluators observed and coded teacher behaviors during classroom visits rather than rating each observation (Stacey et al., 1989).

Training in the use of TPAI was an important component of the implementation of the Teacher Performance Appraisal System (TPAS). Initially, a 30-hour in-service training on effective teaching practices was provided to both evaluators and evaluatees. Evaluators participated in an additional 24-hour course focused on observation and evaluation techniques and a 6-hour course detailing the creation of professional growth

plans. In 1987 and 1988, a 10-hour booster training was provided for evaluators (Stacey et al., 1989).

The recommended evaluation schedule of TPAI was designed for formative evaluations to be conducted minimally three times per year with sufficient intervals of time between the observations to allow for professional development and instructional improvements (Stacey et al., 1989). However, Stacey et al. (1989) pointed out two issues weakened the effectiveness of the formative functions of TPAS. First, the training necessary for evaluators to support teacher development during these intervals was much less extensive than training on observation protocols. Secondly, the connection between areas identified in need of improvement during formative observations and the professional development plans of teachers was overlooked in the appraisal process (Stacey et al., 1989).

As required by North Carolina law, principals are responsible for teacher evaluations; but due to the need for multiple observations of teachers during this new appraisal process, principals received assistance from other trained observers to complete this task (Stacey et al., 1989). While various evaluators reported on the first five functions of TPAI, the principal evaluated the last three that were connected to school operations. A year-end summative evaluation was typically delivered by the principal (Stacey et al., 1989).

In the 1985-1986 school year, the North Carolina TPAS was field tested in 40 school districts. A 1987 study was conducted by NCDPI to evaluate the application of TPAS to a merit pay system and to determine the professional growth and development potential of the system (Stacey et al., 1989). Through a "comprehensive and elaborate survey" (Stacey et al., 1989, p. 85), evaluators and evaluatees responded to their

perceptions of the effectiveness of the structure, process, and outcomes of the system. Both the eight functions of the instrument and the five- or six-point rating scale were considered in the survey, and respondents were asked to rate 10 potential outcomes or benefits from TPAS that impacted teachers, principals, and/or evaluators. The experience levels of evaluators and teachers were considered in the study, and the grade levels that teachers were assigned to were identified. Ten benefits of TPAS were rated on a scale of 1 (not at all) to 3 (significantly beneficial). The 10 benefits were (1) better teaching, (2) teacher job satisfaction, (3) teacher commitment, (4) better staff development, (5) improvement guidelines, (6) easier appraisals, (7) easier employment decisions, (8) more objective appraisals, (9) better communication between evaluator and teacher, and (10) better teacher conferencing.

Stacey et al. (1989) found that evaluators endorsed all 10 benefits of TPAS more highly than did teachers. Whereas, evaluators identified eight of the 10 as a moderate or significant benefit, teachers only rated better teaching as strongly. Middle school teachers endorsed better teaching more than either elementary or high school teachers. The least endorsed TPAS benefit by teachers was job satisfaction, with the most experienced teachers (more than 10 years) being the least likely to endorse job satisfaction. Elementary and middle school teachers endorsed job satisfaction more than high school teachers. Additionally, specific elements of TPAS including postobservation conferencing, postconference follow-up, and the summative evaluation were found to have lower satisfaction ratings among teachers than evaluators (Stacey et al., 1989).

Generally, results from the survey showed that overall TPAS were positively perceived, although the formative functions received more divergent ratings than the summative functions. Stacey et al. (1989) speculated that this was due to the lack of

intensive training in the formative functions of the process. Evaluators rated all of the 10 outcomes of TPAS as more beneficial than did teachers, with only improved teaching rated as highly by teachers as it was by evaluators (Stacey et al., 1989). In actuality, however, respondents identified serious deficits in how the evaluation process promoted professional growth opportunities (Stacey et al., 1989). Other notable findings from the study revealed that the evaluation criteria were acceptable, TPAS improved understandings of instructional goals and effective teaching strategies, and that both teachers and evaluators favored a five- or six-point rating scale (Stacey et al., 1989). Based on these findings, which were supported by a third-party review, TPAS was validated as a workable evaluation system (Bradshaw, 2002).

TPAS studies. When the North Carolina General Assembly enacted the Excellent Schools Act of 1997, TPAS was the preferred evaluation model (Bradshaw, 2002). Although local alternative evaluation systems were acceptable, they were required to undergo a screening process that included the replication of validation studies used with TPAS; thus, districts that had alternatives to TPAS in place were discouraged from continuing to pursue these models (Bradshaw, 2002). Legislators, who were concerned with improving student achievement and reducing teacher attrition, mandated new, tougher standards for teaching performance, increased salary schedules for teachers who had achieved advanced certifications such as National Board status, and proposed a "revised, uniform evaluation system based on student achievement, skills and knowledge, and professional development" (Bradshaw, 2002, p. 116). The result of these political reforms prompted a reexamination of the effectiveness of TPAS (Bradshaw, 2002).

While TPAS generated positive educational changes such as clarified expectations for teachers, the establishment of a common language for discussing

instructional strategies, and a continuing dialogue about instruction among educators, concerns about TPAS were still identifiable (Bradshaw, 2002). According to Bradshaw (1996), teachers and administrators reported numerous issues with TPAS:

1. Lack of a complete list of effective/desirable teaching behaviors;

2. Time-consuming process;

3. Difficulty in meeting both summative and formative purposes;

4. Reduced monitoring of the system due to lack of resources;

5. Inconsistent practices resulting from reduced training and support over time; and

6. Need to link teacher evaluation more directly to student learning.

Citing that only four validation studies of TPAS had been completed since the 1985 implementation and that none of these studies reported current information, Bradshaw (1996) questioned whether the evaluation system was, in fact, valid when the Excellent Schools Act of 1997 was passed.

Referencing *The Personnel Evaluation Standards* (Joint Committee on Standards for Educational Evaluation, 1988), Bradshaw (2002) explained that these standards were created as a means of assessing educational evaluations and relied on four categories to define sound, fair assessment practices and to answer questions of propriety, utility, feasibility, and accuracy in evaluation. Bradshaw (2002) concisely stated, "In a broad sense, a valid teacher evaluation system evaluates what it is designed to evaluate, accomplishes the purpose of the evaluation system, operates fairly, and supports the goals of the organization" (p. 116).

Building on these standards, Bradshaw (2002) conducted an updated study in 1999 of the experiences of teachers and administrators with the TPAS and "their

perceptions of the relationship between teacher evaluation, student learning, teacher development, and school improvement" (p. 118). To gather information, Bradshaw utilized the TEP Questionnaire designed by Stiggins and Duke (1988), including modifications and additional questions that were appropriate to North Carolina. Educators, including 4,092 teachers and 177 evaluators in 21 northeastern school districts, completed the survey with a return rate above 60% (Bradshaw, 2002).

Results from Bradshaw's 1999 TEP indicated that the effectiveness of the TPAS was inconsistent and improvements to the TPAS were recommended (Bradshaw, 2002). For example, Bradshaw (2002) reported that "only 50.4 percent of the teachers and 49.1 percent of the evaluators responded positively" (p.118) in response to items about the quality of the teacher evaluation process. Even fewer evaluators (45.8%) and teachers (32.8%) felt that teacher evaluation positively impacted student achievement. Bradshaw (2002) noted that TPAS was designed to meet both formative and summative purposes; however, 65.7% of evaluators identified the purpose of TPAS as promoting teacher growth. Bradshaw (2002) speculated that "this finding could have reflected a recognition that the results were seldom used for teacher dismissal" (p. 118). Teacher responses on TEP supported this conclusion as 89.5% of those surveyed reported receiving performance ratings that were above standard or consistently superior.

Bradshaw (2002) noted concerns about the validity, reliability, and integrity of TPAS, referencing teacher responses to TEP. For instance, 27% of teachers reported that although one, annual classroom observation was required, they were never observed. Nonetheless, 74.6% of teachers indicated that they trusted their evaluators and believed that the evaluator was credible (68.9%). Fewer teachers (49.5%) felt that the evaluation process was helpful, and approximately 45% reported concerns about limited

professional development resources.

Bradshaw's (2002) research also revealed that in 1999, "less than half of the teachers responding felt that the purpose of teacher evaluation was clear" (p. 121) and that about one-third of the respondents indicated a mid-range response between teacher growth and accountability on TEP. In response to the impact that the teacher evaluation process had on their teaching effectiveness, 54% of teachers reported a positive impact (Bradshaw, 2002). Furthermore, Bradshaw (2002) stated,

> Only a small portion of the teachers reported that teacher evaluation had little or no positive effect on their attitudes about teaching (13.3 percent) and their teaching behaviors and strategies (14.5 percent). On the other hand, almost as many of the remaining teachers gave moderate responses as positive ones. Only 54 percent felt that the evaluation process has a positive impact on the way they felt about themselves as a teacher, and only 46.7% reported a positive impact on their teaching behaviors and strategies. (p. 121)

Based on these findings, Bradshaw (2002) noted that "there was much room for improvement [to the TPAS] and that [the survey results] raised questions about the effectiveness of state mandates" (p. 123).

TPAI-2000. The end of the 1990s saw a resurgence of activity from North Carolina politicians to ensure that the state continued to lead the nation in educational progress. Sparked into action by Governor Hunt's State of the State Address on February 1, 1999, the Education Cabinet created five new goals for education in North Carolina: (1) high student performance; (2) every child ready to learn; (3) safe, orderly, and caring schools; (4) quality teachers and administrators; and (5) strong family, business, and community support. These goals were designed to provide guidance to all education

initiatives with the purpose of building the best public school system in America by 2010 (North Carolina State Board of Education, n.d.).

Revisions to TPAS followed this flurry of political activity (Flowers, Testerman, Hancock, & Algozzine, 2000). Changes to the system were based on legislation that required all certified employees to be evaluated annually with an instrument that incorporated the North Carolina Teacher Evaluation Standards and included improving student achievement, employee skills, and knowledge as well as identifying unsatisfactory and below standard teacher performance (Flowers et al., 2000). The resulting evaluation instrument, TPAI-2000, was designed as a summative tool and was promoted as having the ability to defend against "recent attacks from legislators and the general public regarding the poor performance of some teachers" (Flowers et al., 2000, p. 2) by providing quality assurance that a teacher demonstrated "at least minimal levels of competency" (Flowers et al., 2000, p. 2). Work on TPAI-2000 was the collaborative effort of stakeholders from numerous educational institutes and school districts (Flowers et al., 2000).

Even though the original TPAS, which had been implemented in 1985, was in need of improvements (Bradshaw, 2002), GS 115C-333 and GS 115C-335 provided the impetus for revisions to the system (Flowers et al., 2000). GS 115C-333 gave specific directions as to the number of evaluations for certified employees, using career and school status as criteria for observations and designating actions to be taken if an employee's performance was unsatisfactory or below standard (Flowers et al., 2000). GS 115-335 authorized the use of the North Carolina Teacher Evaluation Standards in evaluation programs and required that performance standards include improving student achievement and employee skills and knowledge (Flowers et al., 2000). Evaluation

models used by other states were considered to meet the demands of this legislation.

However, an advisory committee of teachers and administrators suggested that revisions

of the original TPAI be made, explaining that the TPAI could identify low-performing

teachers and that the adoption of a new evaluation system would require extensive

training for all stakeholders (Flowers et al., 2000). Seven modifications were made to the

original TPAI:

    (1) Indicators were modified to reflect current research and teaching standards.
The indicators for the beginning teacher TPAI were aligned with the TPAI-2000.

    (2) The rating scale was adjusted from a 6-point scale to a 4-point scale and was
the same as used for the beginning teacher TPAI.

    (3) Observation time was decreased by incorporating a Snapshot version of the
observation when appropriate.

    (4) Pre and postconferences were not required if the observation was deemed at or
above standard.

    (5) Preconference interview protocol required the teacher to provide evidence of
teacher practices that include differentiation and alignment of the instruction
to the standards.

    (6) The Formal Observation Data Analysis (FODA) became unnecessary if a
teacher has a history of "at or above standard" ratings.

    (7) Data from the Individual Growth Plans was reviewed and evaluated for
alignment to local and state goals (Flowers et al., 2000).

Appendix B details TPAI-2000.

One unique aspect of TPAI-2000 was the separation of the summative and

formative aspects of evaluation (Flowers et al., 2000). Although professional growth was given as one of the two purposes of this evaluation process (with the other being accountability and quality assurance), it was viewed through the lens of a summative instrument. In the "Philosophy and Purpose," the authors stated, "An effective evaluation program results when teachers are treated as professionals as well as when evaluators are successful in using evaluation to reinforce effective practices and to improve teaching" (Flowers et al., 2000, p. 1). Systems were "***strongly suggested"*** (Flowers et al., 2000, p. 16) to have "a structured formative evaluation system, also called an alternative evaluation system, in combination with a summative evaluation system" (Flowers et al., 2000, p. 16) and to allow teachers choice in the utilization of formative evaluation strategies.

Flowers et al. (2000) stated that creators of TPAI-2000 explicitly discussed separating the function of summative and formative evaluations, explaining,

> Combining purposes, accountability and professional growth, into one evaluation system is very controversial. Data for accountability are used to make high-stakes decisions concerning the experienced teacher's career. The experienced teacher should put her or his best foot forward and 'show off' the best she or he has to offer. For professional growth, the experienced teacher is asked to take risks and explore new and different techniques. The experienced teacher needs to feel comfortable and know that these data will <u>not</u> be used to make high-stakes decisions. . . . Data collected are either summative or formative and should not be mixed. (p. 17)

Other recommendations set forth by the authors of TPAI-2000 listed strategies that were perceived to build trust between the evaluator and the teacher so the evaluation process

could lead to continuous instructional improvement. These recommendations included

the use of multiple evaluators and data sources, well-trained evaluators, explicit criteria

and procedures for data collection, flexibility when observing various teaching styles, and

clear communication between participants in the process (Flowers et al., 2000).

Twenty-first century reforms in North Carolina. North Carolina's history of

enacting educational reform to ensure student achievement continued into the 21st

century as leaders considered how to best prepare students for the demands of a new

millennium (Public Schools of North Carolina, 2012a). As a result of this revisioning, an

updated mission statement and five supporting goals for future-ready students were

adopted by the State Board of Education in September 2006 (Public Schools of North

Carolina, n.d.b). Appendix C shows the updated mission and goals.

Shortly following the adoption of this new mission statement and supporting

goals, the North Carolina Professional Teaching Standards, which were originally

approved by the State Board of Education in 1998, were redefined to reflect "what

teachers need to know and be able to do in 21st century schools" (Public Schools of

North Carolina, 2013b, n.p.). These new standards were put into place in June 2007

(Public Schools of North Carolina, 2013b). By 2008, the State Board of Education had

amended the teacher evaluation process to align with the North Carolina Professional

Teaching Standards (Public Schools of North Carolina, 2009) and approved the

Framework for Change: the Next Generation of Assessment and Accountability (Public

Schools of North Carolina, 2012a). Packaged as the READY initiative, these changes

also incorporated a new standard course of study that included Common Core State

Standards, new assessments, and a refashioned accountability model to replace the state's

first accountability system, the ABCs of Public Education, which was implemented in the

1996-1997 school year (Public Schools of North Carolina, 2012a).

Most READY initiatives were scheduled to be operational in the 2012-2013 school year (Public Schools of North Carolina, 2012a). However, NCTEP was implemented in three phases between 2008-2009 and 2010-2011 with the expectation that all teachers would be evaluated with the new instrument in the last phase of the rollout (Public Schools of North Carolina, 2010). The 2008 version of NCTEP was based on the *Framework for 21st Century Learning* and the five North Carolina Professional Teacher Standards:

(1) Teachers demonstrate leadership.

(2) Teachers establish a respectful environment for a diverse population.

(3) Teachers know the content they teach.

(4) Teachers facilitate learning for their students.

(5) Teachers reflect on their practice (Public Schools of North Carolina, 2009).

The process for completion of NCTEP was outlined in the North Carolina State Board of Education Policy TCP-C-004, dated October 2, 2008. Components of the process included training, orientation, self-assessment, pre and postconferences and observations, observations based on years of experience and certification levels, a summative evaluation conference, and professional development plans that were contingent on performance ratings (Public Schools of North Carolina, 2009). The Rubric for Evaluating Teachers and a Summary Rating Form allowed for teacher performance to be described as developing, proficient, accomplished, or distinguished or not demonstrated. If not demonstrated was selected, a commentary by the observer was required. Directions for completing the rubric and summative form were given in the NCTEP manual and were available to both teachers and evaluators (Public Schools of

North Carolina, 2009). The evaluation instruments were developed in conjunction with the Mid-continent Research for Education and Learning (McREL), a non-profit organization whose purpose was to improve education through applied research, product development, and service (Williams, 2009).

North Carolina teacher evaluation and RttT reforms. Although North Carolina educational reforms were well underway by 2010, RttT grant funding accelerated initiatives in the targeted areas, also designated as pillars, of

- Great Teachers and Principals

- Learning Standards and Assessments that Align to 21st Century Demands

- Data and Technology Systems to Support Schools

- Turnaround Support for Lowest-Achieving Schools (NCDPI, 2015).

Under the pillar of Great Principals and Leaders, teacher and principal evaluation was identified as an area to be advanced by "developing an equitable, reliable, valid, and transparent evaluation that will inform professional development or staffing needs, and identify highly effective teachers" (NCLEG, 2010, p. 3). The North Carolina RttT Evaluation Plan articulated three key endeavors necessary for progress in this area:

1. Adding a student growth component to the Educator Evaluation process;

2. Fully implementing an Educator Evaluation process that includes student growth measures as essential and significant components of both the Teacher Evaluation Process and the Principal Evaluation Process; and

3. Conducting a thorough, data-informed planning and evaluation process − via a *Teacher Effectiveness Workgroup* − with all relevant constituents represented to (a) determine the most rigorous, transparent, and fair way to incorporate

student growth measures in all teacher and principal evaluations and (b) study

teacher compensation models in place across the state and evaluate their

fairness, validity, and reliability (Public Schools of North Carolina, 2010).

Due to the RttT funding, North Carolina made great progress toward these goals by

2012-2013 (NCDPI, 2015). A notable indicator of these efforts was the inclusion of a

sixth standard, Teachers Contribute to the Academic Success of Students, in the NCTEP

manual revisions of September 2012 (Public Schools of North Carolina, 2012b). Teacher

ratings under Standard 6 relied on measures of student growth as calculated through the

Educational Value-Added Assessment System (EVAAS; Henry & Guthrie, 2015).

Further evidence of progress was the establishment of online tools to support the

professional development needs and to facilitate data collection during the evaluation

process (Public Schools of North Carolina, n.d.d).

RttT Reporting on Educator Evaluation Initiatives

Throughout the RttT grant funding period, periodic status reports were required

by the United States Department of Education to determine the impact and effectiveness

of North Carolina's objectives (NCDPI, 2015). The importance of these evaluations was

communicated in the *North Carolina Race to the Top Closeout Report: Executive*

Summary (2015):

From the beginning of RttT, the State has been committed to thorough,

independent, and rigorous policy and program evaluations. To that end, the State

contracted with a consortium of North Carolina universities to provide on-going

formative and summative evaluations of each RttT initiative. The evaluation

reports, delivered at least annually, have helped to guide implementation of each

initiative by providing data and feedback regarding fidelity to plan and, where

feasible, early indicators of effectiveness. This information has been critical to

informing mid-course corrections and design and management of new programs,

and has also provided a body of knowledge that is accessible to the public as well

as policy makers across the country. As RttT approaches its end, these evaluation

reports are shifting to summative evaluations that intend to assess programs' "cost

versus benefit" and cost-effectiveness. This information will be valuable in

informing future funding, policy, and management decisions. As of January 23,

2015, the NC RttT Evaluation initiative has released 51 reports; six more will be

provided by the end of June, 2015. (NCDPI, 2015, p. 9)

All evaluation reports for RttT were conducted by the Consortium for Educational

Research and Evaluation-North Carolina (CERE-NC) and are accessible at

http://cerenc.org/rttt-evaluation/executive -summaries/ (NCDPI, 2015).

Evaluation reports revealed glimpses of teacher perceptions about the state of

educator evaluation in North Carolina (Davis et al., 2015) and gave insight into the

feasibility (Rose, Henry, & Lauen, 2012) and impact (Henry & Guthrie, 2015) of teacher

effectiveness measures on teacher ratings as stipulated by the RttT initiatives (Public

Schools of North Carolina, 2010). In an early study entitled "Comparing Value-Added

Models for Estimating Individual Teacher Effects on a Statewide Basis: Simulations and

Empirical Analyses," researchers examined eight value-added models and identified four

of these models as appropriate for low-stakes purposes, noting that none of the four

models "performed sufficiently well to be considered for high stakes purposes" (Rose et

al., 2012). EVAAS was one of the four acceptable models (Rose et al., 2012) and was

selected by the State Board of Education in 2012 as the North Carolina model used to

measure student growth (Lynn et al., 2013).

Lynne et al. (2013). Additional reporting on student growth in NCEES examined the inclusion of Standard 6 in NCTEP (Lynn et al., 2013). The Standard 6 rating was based on EVAAS measures to determine an educator's impact on student growth, and student growth was defined by the change in students' standardized test performance over time (Lynn et al., 2013). According to evaluators, Standard 1 through Standard 5 were given the same weight as Standard 6 when the overall effectiveness status of an educator was assessed (Lynn et al., 2013).

A timeline was set by the State Board of Education in April 2013 that after 3 years of data collection, educators would receive an overall status rating as well as the annual ratings for each standard (Lynn et al., 2013). The 2014-2015 school year was approved by the U.S. Department of Education as the first year for statuses to be provided (Lynn et al., 2013). Since EVAAS measures were based on standardized test scores, and EOG/EOC tests were unavailable for over 60% of North Carolina teachers, alternative standardized assessments were created to allow growth measures to be calculated for all teachers (Lynn et al., 2013). Measures of student learning were planned to provide these measures for all subjects and grades (Lynn et al., 2013). From 2011 to 2013, several adjustments to the calculation of teacher effectiveness were made due to issues with the impact of inclusion of school-wide growth measures on individual teacher ratings (Lynn et al., 2013). Also, in this time period, RttT Professional Development, focusing on NCEES, was provided by NCDPI (Lynn et al., 2013).

In addition to providing background information about the expansion of NCEES to include student growth, the formative evaluation report "Measures of Student Growth in the North Carolina Educator Evaluation System" addressed the correlations of Standard 6 to the other five standards, the use of growth measures to inform instruction

and evaluation, and described educator perceptions of the use of growth data in evaluation (Lynn et al., 2013). Findings revealed that

> Teachers' views of their own efficacy and of the degree to which they believe they have prepared their students are significantly and positively related to their value-added measure, while their sense of fairness of the evaluation process has no statistical relationship with their value-added measure. (Lynn et al., 2013, p. 4)

The correlation between student growth and other teacher evaluation measures by subgroups was limited due to the majority of teacher scores falling in the proficient and accomplished categories (Lynn et al., 2013). As to the use of data to inform instruction and evaluation, results were mixed, showing that not all educators were fully using EVAAS data (Lynn et al., 2013). Teacher perceptions of the use of growth data in evaluation revealed that uncertainties and misconceptions about the calculation of effectiveness ratings impeded the use of growth data as a tool for informing instruction (Lynn et al., 2013). Publication of "They Just Want to Fire Me! NC's Educator Effectiveness Myths Debunked" on NCDPI's Educator Effectiveness website addressed some of these misconceptions (Public Schools of North Carolina, 2014).

Four recommendations resulted from the Lynn et al. (2013) study. These recommendations were to (1) improve communication, (2) expand training about Standard 6, (3) increase opportunities for teacher input, and (4) to make revisions of the NCTEP ratings and evaluation process (Lynn et al., 2013). With regard to the need for revision of NCEES, analysts noted concerns about inflated subjective ratings by evaluators on Standards 1 to 5 ratings, referencing the Widget Effect (Weisberg et al., 2009) that occurs when evaluator ratings are "upwardly biased or benchmarked to

minimum requirements when they are used in summative evaluation" (Lynn et al., 2013, p.6). Weisberg et al. (2009) contended that the existence of an underlying cultural mindset in schools assumes that teachers are synonymous to interchangeable parts. According to Weisberg et al., both teacher expectations and observer perceptions of teacher evaluations are skewed due to this deep rooted attitude, and they point to their research as an illustration of "a policy framework that rarely considers teacher effectiveness for key decisions" (p. 4) that is descriptive of "schools [that] are indifferent to instructional effectiveness–except when it comes time to remove a teacher" (p. 4). To counter rater bias that results from the Widget Effect, Lynn et al. (2013) recommended expansion of rating scales and improved evaluator training.

Henry and Guthrie (2015). Henry and Guthrie (2015) stated, "One of the most important purposes of teachers' evaluations is to identify teachers who need improvement so leadership can intervene in ways that help ensure that students have access to high-quality teaching" (p. 2); however, their report demonstrated, "Value-added scores, including but not limited to EVAAS scores, provide objective measures of the outcomes of teachers' instructional practices but, unfortunately, do not provide information about which practices are strengths and weaknesses for individual teachers" (p. 2).

Considering five research questions concerning outcomes of teacher evaluation since the addition of Standard 6, researchers compiled a database of teacher evaluation ratings and EVAAS scores and merged this database with other information about students, teachers, and schools collected for prior RttT evaluations (Henry & Guthrie, 2015). Also, research from the RttT Omnibus Teacher and Principal Survey was used to identify teacher perceptions about NCEES (Henry & Guthrie, 2015). Sample RttT Omnibus Survey Dimensions and Items, adapted from Henry, Campbell, Thompson, and

Townsend (2014) are included in Appendix D. Other data included observation data

from the Classroom Assessment Scoring System (CLASS) that was used to

 "describe perspectives of teachers and principals on the measure of student achievement

growth and classrooms and to analyze instructional practices associated with increases in

EVAAS scores" (Henry & Guthrie, 2015, p. 7). Finally, data from the Tripod student

survey, which had been piloted in 38 North Carolina school districts in 2011-2012, was

incorporated into the evaluation data (Henry & Guthrie, 2015).

Findings from the study revealed that EVAAS measures were the primary

identifier of teachers who were assigned to the needs improvement categories of not

demonstrated or developing as principal ratings on Standards 1 through 5 were most

often proficient or higher (Henry & Guthrie, 2015). This global rating trend by principals

contributed to less refined feedback to teachers about their strengths and weaknesses and

decreased the credibility of NCEES since evaluation growth measures only loosely

correlated to principal ratings (Henry & Guthrie, 2015). Teacher interview responses

indicated their desire for more and better feedback from evaluators (Henry & Guthrie,

2015).

Additional findings gave insight about instructional practices that contributed to

improved EVAAS scores (Henry & Guthrie, 2015). These practices were determined

through examination of principal ratings, student surveys, and classroom observations

(Henry & Guthrie, 2015). Three categories of practices emerged that impacted value-

added scores: (1) effective practices, (2) practices associated with teacher capacity to

improve, and (3) ineffective practices (Henry & Guthrie, 2015). Table 3 details these

practices, and the sources indicated an impact on EVAAS scores. Researchers stipulated

that discussions of these measures should be included in the feedback provided to

teachers (Henry & Guthrie, 2015).

Table 3

Measures Associated with Impacting Teachers' Value-Added Scores

Effective Practices	Source
Facilitating Student Learning	Principal Ratings
Collaborative Environment	Student Survey – Elementary
Classroom Management	Student Survey – Elementary
Higher-Order Instruction	Student Survey – Secondary
Classroom Management	Student Survey – Secondary
Positive Climate	Classroom Observation
Ineffective Practices	Source
Busy Work	Student Survey – Secondary
Student-Led Environment	Student Survey – Secondary
Teachers' Capacity to Improve	Source
Reflection on Practice	Principal Ratings
Collaborative Environment	Student Survey – Elementary
Higher-Order Instruction	Student Survey – Secondary
Classroom Management	Student Survey – Secondary

The final question of the Henry and Guthrie (2015) study examined "teachers' views about their evaluations and related topics during the period in which the NCEES evaluation with six standards has been implemented" (p. 7). Analysts reported two specific concerns for investigation that arose surrounding the addition of Standard 6:

1. Would perceptions of the overall fairness of evaluations affect the teacher development elements of the NCEES?

2. Would high-stakes evaluations result in less collaboration among teachers (Henry & Guthrie, 2015)?

Responses from teacher surveys illustrated a significant decline in teacher overall views

about evaluations between 2011-2012 and 2013-2014, changing from a slight agreement with survey items to neither agreement nor disagreement (Henry & Guthrie, 2015). Analysts reported that

> The largest declines in favorability related to the developmental use of the evaluation. For example, the largest decline was on the item "The teacher evaluation process encourages professional growth," closely followed by "The evaluation process encourages teachers to reflect on their instructional practice."
> (Henry & Guthrie, 2015, p. 16)

Notably, the decline in the favorability of teacher views appears to be related to disappointment in the formative aspects of NCEES (Henry & Guthrie, 2015).

On a more positive note, survey responses indicated that teachers were more engaged in collaborative activities and participated in these activities on an almost daily basis (Henry & Guthrie, 2015). This shift reflected the job-embedded, student-focused nature of knowledge sharing by teachers (Henry & Guthrie, 2015). Concerns about the potential negative impact of high-stakes evaluation on teacher willingness to work together seemed to be unsubstantiated by the data (Henry & Guthrie, 2015).

The primary issue discovered in Henry and Guthrie (2015) was that EVAAS was, in essence, the only "systematic data source" useful for determining teacher effectiveness (p. 18). The authors suggested including other sources of data such as "student surveys and systematic, direct observations of classrooms" into the teacher evaluation process (Henry & Guthrie, 2015, p. 17). They proposed that inclusion of multiple data sources could improve teacher perceptions of NCEES and inform more actionable feedback. Additionally, new data sources could assist with more accurate identification of teachers in need of improvement (Henry & Guthrie, 2015).

Davis et al. (2015). The "Teacher and Principal Perceptions of the North Carolina Educator Evaluation System: Final Evaluation Report" was released in September 2015 (Davis et al., 2015). This report addressed two specific research questions.

1. How are teachers and principals using EVAAS data for evaluation purposes and to inform teaching practices?

2. What are teachers' and school leaders' perceptions of the use of growth data in the evaluation? (Davis et al., 2015, p. 3)

The Evaluation Team conducted qualitative interviews to gather data about these questions and also used this information to examine "participants' views of the professional development resources and information they received, as well as the purpose, process, and effectiveness of their evaluations and the NCEES process as a whole" (Davis et al., 2015, p. 8). The team noted that although the overall evaluation of teacher perceptions used a mixed-methods approach, this report focused only on the qualitative components (Davis et al., 2015).

Key areas for consideration emerged from teacher and principal interviews (Davis et al., 2015). Concerning the evaluation process and feedback, teacher knowledge-sharing practices, and data-driven instruction, respondents reported mostly positive outcomes (Davis et al., 2015). NCEES was credited for prompting "deeper and more substantive conversations between teachers" (Davis et al., 2015, p. 13), which encouraged "a change in the school culture that resulted in a communicative environment for educators" (Davis et al., 2015, p. 13). Furthermore, "teachers noted that data pulled from the NCEES and other assessments helped them plan for differentiated instruction, highlight areas of strength and weakness within their curricula, and become more

reflective educators" (Davis et al., 2015, p. 14).

Perceptions of the use of student growth data in NCEES were more problematic. Davis et al. (2015) reported that participants indicated that "more clarity on the overall evaluation tool" (p. 16) was warranted and revealed that Standards 1 through 5 were more clearly understood by teachers than Standard 6 (Davis et al., 2015). Conflicting views about the use of student growth data were evident, with some respondents viewing it as "an unnecessary and biased assessment" (Davis et al., 2015, p. 17). Perceptions of the fairness of NCEES included discussions of factors that were not considered in the evaluation process (Davis et al., 2015). Remarking that NCEES had impacted educator morale by creating additional pressure to perform,

> participants shared three ways to make the instrument fairer: 1) account for extenuating circumstances, such as a student's home environment, behavior, and ability; 2) include a section in one of the standards that evaluates teachers' relationships and interactions with students inside and outside the classroom; and 3) reduce the weight of the student growth standard to less than that of the remaining standards. (Davis et al., 2015, p. 17)

Recommendations gleaned from participant interviews reflected the need to continue gathering teacher input on improving NCEES, providing ongoing observation feedback to teachers, and capitalizing on the use of the NCEES to encourage collaboration (Davis et al., 2015). Expansion of training on Standard 6 was also suggested (Davis et al., 2015). Finally, researchers recommended that a standard that addresses the relational aspects of teaching be developed and added to NCEES (Davis et al., 2015). This consideration was based on teacher and principal reports about the significance of student-teacher relationships and suggested the use of student surveys as a

potential measure of a relationship component (Davis et al., 2015).

Corn, Bryant et al. (2013); Corn, Smart et al. (2013); Corn, Halstead et al. (2012). Effective utilization of NCEES was included in RttT Professional Development initiatives (Corn, Bryant et al., 2013). The Professional Development Teacher Survey showed that 63% of teachers agreed or strongly agreed that their knowledge and skills about the revised NCTEP were enhanced due to these RttT professional development initiatives (Corn, Smart et al., 2013). Items from the Omnibus Teacher and Principal Survey, which was created to "assess change across a wide range of constructs that may be influenced by the collective set of NC RttT activities" (Corn, Halstead et al., 2012, p. 9), generated information about educator perceptions of teacher evaluations (Corn, Smart et al., 2013). The Omnibus Survey specified that in 2012, 51% of teachers agreed or strongly agreed that they were overall satisfied with the teacher evaluation process (Corn, Smart et al., 2013).

Smart et al. (2015). The "Fourth Annual Race to the Top Professional Development Evaluation Report" confirmed the "consistent decline in teachers' perceptions of and satisfaction with the teacher evaluation process" (Smart et al., 2015, p. 52). By 2013-2014, an 11% decrease in teacher perceptions about the positive impact of the evaluation process on professional growth was reported, down to 49% in 2013-2014 as compared to 60% in 2011-2012 (Smart et al., 2015). Likewise, "teacher perceptions of the fairness of the evaluation process also continued to decline, with fewer than half of teachers surveyed in 2013-14 agreeing that the evaluation is fair, down 8 percentage points from 2011-12 (53% to 45%)" (Smart et al., 2015, p. 52). Teacher interviews disclosed

mixed emotions on the impact the new teacher evaluation process [had] on

instruction. Some believed principal feedback would help make them become more reflective practitioners while others thought the evaluations would only shift their attention from professional improvement to "teaching to the test." (Smart et al., 2015, p. 53)

Marks et al. (2015). The "North Carolina Race to the Top: Overall Impact and Implementation Findings – Final Report" assessed how North Carolina's RttT pillars of reform had strengthened North Carolina's workforce (Marks et al., 2015). The effectiveness of the new NCEES process was considered under Pillar 2: Great Teachers and Leaders. Of note, "North Carolina developed more initiatives in support of this pillar than for all other pillars combined in an effort to provide multiple routes to enhance the states' capacity to increase the effectiveness and equitable distribution of educators" (Marks et al., 2015, p. 28). To determine the enhanced capacity of educators,

the [Evaluation] Team documented "the extent to which North Carolina's RttT funding supported growth in educators' capacity to make data-driven decisions, provide and support great teachers and leaders, turn around the lowest achieving schools, implement high standards, and use assessments effectively for formative, benchmark, and summative purposes." (Marks et al., 2015, p. 15)

Unfortunately, data collected between 2010-2011 and 2013-2014 did not demonstrate that principal ratings on Standards 1 through 5 denoted increased levels of higher-quality teaching (Marks et al., 2015). Likewise, first-year teacher ratings showed no significant change across the years studied. Teacher value-added scores were indicative of this trend. No statistically significant findings demonstrated increased effectiveness when comparing changes in performance to the rate of improvement prior to RttT (Marks et al., 2015). With respect to teacher perceptions about teacher

evaluations, ratings as to the fairness of the North Carolina Evaluation System decreased

between 2012 and 2014 (Marks et al., 2015).

Marks et al. (2015) remarked that "determining the impact of a multi-year, state-level

intervention comprised of multiple initiatives in multiple locations with a variety of

implementers is challenging at best" (p. 32). They noted important limitations to the

study, including

1. Difficulty in establishing a valid comparison group against which to measure

   initiatives;

2. Difficulty in determining the true impact of RttT initiatives since other

   services may also have affected outcomes;

3. Difficulty in discerning the impact of changes in leadership and economic

   conditions on RttT outcomes that were independent of the initiatives

   themselves; and

4. The limited timeframe from "initial implementation to outcome

   measurement[that] may not yet have been sufficient to capture many of the

   initiatives' eventual intended effects." (Marks et al., 2015, p. 32)

Over the course of the RttT project, 19 amendments were made to the original

RttT Scope of Work. Thirteen of these amendments were made in connection to the

Great Teachers and Leaders Pillar (Marks et al., 2015). The number of amendments

indicate the "complexity of managing and implementing a plan on the scale of RttT" over

an extended period of time (Marks et al., 2015, p. 37).

In conclusion, the "North Carolina Race to the Top: Overall Impact and

Implementation Findings – Final Report" emphasized the impact of the limitations on the

analysis of RttT outcomes, especially in the areas of teacher effectiveness and student

achievement. The authors postulated, "While there was no evidence of RttT impact on either teacher effectiveness or most measures of student achievement, without a valid comparison group it is not possible to determine whether this lack of evidence reflects meaningful performance outcomes" (Marks et al., 2015, p. 33). Regardless of this uncertainty, the authors continued by stating, "Despite these limitations, however, the data can and do support the conclusion that the education landscape in North Carolina is indeed better off at the conclusion of RttT than it was before" (Marks et al., 2015, p. 33).

Connected studies. While RttT studies provided information about the impact of the changes to NCTEP on teacher perceptions of the process, especially as it related to the addition of a student growth measure (Henry & Guthrie, 2015; Lynn et al., 2013), an exploratory study by Leggett (2015) utilized the North Carolina teacher evaluation instrument to examine "the influence of principal and school characteristics on principal ratings of teachers" (p. 147). Leggett based her study on Bandura's (1986) social cognitive theory and the "premise that human behavior (principal ratings of teacher) cannot be separated from personal characteristics of the individuals involved (the principals) and the environmental characteristics of the context (the school)" (p. 147). Leggett's study concluded that factors such as principal years of experience, teacher attrition rate at the school, and the number of beginning teachers in a school were predictive of variances in principal ratings of teachers. Implications from Leggett's research indicated that more study of principal characteristics in relation to their ratings of teachers is necessary and that policymakers should consider all factors that influence principal ratings when evaluation processes are being developed or revised.

NCTEP (2015)

Teacher evaluation, as outlined by the NCTEP manual (2015), is a

comprehensive, performance-based model (Public Schools of North Carolina, 2015).

This manual identifies the purpose and goals of the evaluation process, defines the North

Carolina Professional Teaching Standards, explains the eight components of NCTEP, and

provides examples of evaluation rubrics and related forms. As reported on NCDPI's

Educator Effectiveness Model webpage, an electronic data collection system, NCEES is

used during NCTEP (Public Schools of North Carolina, n.d.a). NCEES houses all

evaluation information on a web-based platform, hosted by True North Logic (Lynn et

al., 2013).

The North Carolina Professional Teaching Standards serve as the basis for

NCTEP and also underpin teacher preparation and professional development (Public

Schools of North Carolina, 2013b). In Standards 1 through 5, the skills and knowledge

teachers are expected to possess are defined (Public Schools of North Carolina, 2015).

Standard 6, which was added to NCTEP by the North Carolina State Board of Education

in February 2012, simply stated, "The work of the teacher results in acceptable,

measurable progress for students based on established performance expectations using

appropriate data to demonstrate growth" (Public Schools of North Carolina, 2015, p. 12).

Whereas the evaluation of Standards 1 through 5 is completed through the work of

observers, Standard 6 is based on "a student growth value as calculated by the statewide

growth model for educator effectiveness" (Public Schools of North Carolina, 2015, p.

40). In North Carolina, EVAAS, a statistical analysis system, calculates student growth

rates and determines teacher effectiveness that is included in Standard 6 (Lynn et al.,

2013).

According to the NCTEP manual (2015), "The intended purpose of the North

Carolina Teacher Evaluation Process is to assess the teacher's performance in relation to

the North Carolina Professional Teaching Standards and to design a plan for professional

growth" (Public Schools of North Carolina, 2015, p. 4). This purpose is further defined

by eight goals as follows:

The teacher performance evaluation process will:

- Serve as a measurement of performance for individual teachers;

- Serve as a guide for teachers as they reflect upon and improve their

  effectiveness;

- Serve as the basis for instructional improvement;

- Focus on the goals and objectives of schools and districts as they support,

  monitor, and evaluate their teachers;

- Guide professional development programs for teachers;

- Serve as a tool in developing coaching and mentoring programs for teachers;

- Enhance the implementation of the approved curriculum; and

- Inform higher education institutions as they develop the content and

  requirements for teacher training programs. (Public Schools of North

  Carolina, 2015, p. 5)

To meet these goals, NCTEP incorporates a performance-based evaluation model

(Shakman et al., 2012). This comprehensive model includes eight components: (1)

training, (2) orientation, (3) self-assessment, (4) preobservations, (5) observations, (6)

postconferencing, (7) a summary evaluation conference conducted by the principal

including a summary rating of performance, and (8) professional development plans

(Public Schools of North Carolina, 2015). The Figure below illustrates NCTEP. As

required by G.S. 115C-333.1(a), all teachers are evaluated at least annually, and teachers

who demonstrate unsatisfactory performance may be placed on mandatory improvement plans. Annual evaluations are based on one of three evaluation cycles: comprehensive, standard, or abbreviated (Public Schools of North Carolina, 2015).

COMPONENT 2: Orientation

Within two weeks of teacher's first day, the principal will provide:
A. The Rubric for Evaluating North Carolina Teachers;
B. Teacher Evaluation Policy **ID Number: TCP-C-004**; and
C. A schedule for completing evaluation process.

COMPONENT 3:
Teacher Self-Assessment

Using the Rubric for Evaluating NC Teachers, the teacher shall rate his or her performance and reflect on his or her performance throughout the year.

COMPONENT 1:
Training

Before participating in the evaluation process, all teachers, principals and peer evaluators must complete training on the evaluation process.

COMPONENT 4: Pre-Observation Conference

Goal: To prepare principal for the observation. Before the first formal observation, the principal meets with the teacher to discuss: **self-assessment, professional growth plan and a written description of the lesson(s)** to be observed.

STEP 1:
Training and Orientation

STEP 2:
Self-Assessment, Goal Setting and Pre-Conference

STEP 4:
Summary Evaluation and Goal Setting

STEP 3:
Observation Cycle (Administrative and Peer)

COMPONENT 8:
PD Plans

A. **Individual Growth Plans:** "Proficient" or better
B. **Monitored Growth Plans:** at least 1 "Developing"
C. **Directed Growth Plans:** "Not Demonstrated" or "Developing" rating for 2 sequential years

COMPONENT 5:
Observations

A. *Formal observation:* 45 min. or entire class period. *Informal:* at least 20 min.

B. **Teachers <3 years employment:** *Comprehensive Evaluation Cycle:* 3 formal (principal) & 1 formal (peer)

C. **Teachers >3 years employment:** *Standard or Abbreviated Evaluation Cycle:*

Standard: 3 Observations, 1 must be formal

Abbreviated: 2 Observations on Standards 1 and 4. Abbreviated Observations may be formal or informal.

Teachers Renewing License: *Standard Evaluation Cycle*

COMPONENT 7:
Summary Evaluation Conference and Scoring the Teacher Summary Rating Form

Prior to end of school, the principal conducts a summary evaluation conference with teacher to discuss components of the evaluation cycle type used: Comprehensive or Standard. At the conclusion:
A. Give rating for each Element in Rubric for Evaluating NC Teachers;
B. Comment on "Not Demonstrated;"
C. Give an overall rating of each standard observed;
D. Provide teacher with opportunity to add comments to the Summary Rating Form;
E. Review completed Teacher Summary Rating Form with teacher; and
F. Secure the teacher's signature on the Record of Teacher Evaluation Activities and Teacher Summary Rating Form.

COMPONENT 6:
Post-Observation Conference

The principal shall conduct a post-observation conference no later than ten school days after each formal observation.

Discuss and document strengths and weaknesses on the Rubric.

Figure. Teacher Evaluation Process Infographic May 2015
(http://www.ncpublicschools.org/docs/effectiveness-model/ncees/instruments/teach-eval-manual.pdf)

Summary

The extent of literature on current practices in teacher evaluation spans nearly 4 decades (Darling-Hammond, 2014). While studies such as the Measures of Teacher Effectiveness (MET) research demonstrate the significance of teacher quality on student achievement (MET Project, 2013) and new evaluation systems have been implemented nationwide in response to RttT initiatives (Mead, 2012), teacher perceptions of these innovations have only recently begun to be investigated (Donaldson, 2012). Since researchers propose that certain minimal conditions must be present for an effective educator evaluation to exist (Darling-Hammond et al., 1983) and that certain attributes of the evaluation process contribute to the quality and impact of the evaluation experience (Stiggins & Duke, 1988), further study of teacher perceptions of the impact of NCTEP on their teacher effectiveness, their professional growth, and their attitudes toward teaching is warranted.

Chapter 3: Methodology

Introduction

Since it has long been recognized that educator evaluation has the potential to enhance professional growth and teacher effectiveness (Stiggins & Duke, 1988), it is important to understand how teachers perceive the evaluation process as it relates to their teaching practice (Lee, 2011). Researchers at CERE-NC have reported that teacher attitudes toward NCEES became less favorable between 2011-2012 and 2013-2014 (Henry & Guthrie, 2015). Although concerns existed about reforms that included VAMs in the evaluation process, it appears that the less favorable attitudes "reflect disappointment in the developmental purposes for which the evaluations were conducted, rather than in fairness or the specific standards" (Henry & Guthrie, 2015, p. 16). The purpose of this study was to examine educator perceptions of the impact NCTEP has on professional growth and teacher effectiveness as well as how the evaluation process impacts their attitudes toward teaching.

Participants

Participants were teachers of Grades PreK-13 students from a midsized school district located in the foothills of North Carolina. Permission for teacher participation in this study was granted by the superintendent. Principals were asked to endorse the study. Participants of this study worked in one or more of the system's 26 schools, and some participants worked with multiple grade levels. Grade-level configurations in the district were 11 PreK-5 elementary schools; four 6-8 middle schools; four K-8 schools; three high schools serving Grades 9-12; an Early College High School and a Middle College High School serving Grades 9-13; and two alternative schools, one that serves Grades K-5 and a second that serves Grades 6-12. The Middle College High School and Early

College High School were located on the campus of the local community college; however, teachers at these sites were employed through the county's school system.

For the purposes of this study, the grade level(s) teachers were assigned to was considered; however, the specific school(s) was not identified. Other participant demographics that were collected included the total years of teaching experience; grade level taught at the time of the most recent, completed annual evaluation cycle; and teacher status rating in this same evaluation cycle. Survey answers remained anonymous unless participants choose to provide a contact number or email for a follow-up interview with the researcher. If this was the case, a coding system was used to identify participants who were involved in the interview phase of this study. Participation in this study was voluntary.

Instruments

TEP modified. A modified version of TEP was used to gather data from all teachers in the district at the end of the 2015-2016 school year. Permission for use of the modified version of TEP was obtained from Dr. Daniel Duke, one of the original authors of the instrument. Appendix E contains email correspondence that documents Dr. Duke's consent. Through the original TEP, Stiggins and Duke (1988) validated the five key attributes of effective evaluations and utilized the instrument to determine if these attributes were related to perceived teacher growth and improved teacher effectiveness. In this researcher's study, the five key attributes of effective evaluation were examined in relation to four outcome rating questions. Attribute categories were (1) attributes of the teacher, (2) attributes of the person who observes and evaluates, (3) attributes of the procedures used to gather data on teacher performance, (4) attributes of the feedback, and (5) attributes of the evaluation context (Stiggins & Duke, 1988). Outcome rating

questions asked for teacher perceptions of the impact of the evaluation process on the overall quality of the evaluation experience and the impact of the evaluation process on teacher understandings of the learning process (teacher effectiveness), teaching strategies and behaviors (professional growth), and attitudes toward teaching.

The original TEP asked teachers to describe 55 attributes within each of the five categories of effective evaluation: teacher, evaluator, procedures, feedback, and context. In addition, teachers were asked to rate the outcome of their last evaluation considering all components of the process and to rate the impact of the experience on three aspects of their professional practices. The first aspect investigated how teacher attitudes were changed about content, students, or themselves as a result of the evaluation. The second aspect considered how teacher behaviors and strategies changed due to the evaluation process. The third aspect examined the impact of the evaluation on teacher perspectives of changes in their effectiveness in relation to the evaluation process (Stiggins & Duke, 1988). Changes to the modified TEP reflect NCTEP but otherwise follow an identical format to the original TEP. Dr. Duke (personal communication, April 12, 2016) reviewed changes to the original TEP and stated that it aligned well with NCTEP and that validation issues should not be of concern. More information about Stiggins's and Duke's work on teacher evaluation as well as the original TEP Questionnaire is located in Stiggins and Duke's (1988) work.

TEP (modified) reflected the same key elements as the original TEP instrument and was designed with a similar purpose: to identify teacher perceptions of the overall evaluation process as it relates to teacher attitudes toward teaching, professional growth (behaviors), and teacher effectiveness (understandings). Revisions to the original TEP's directions and questionnaire items reflect NCTEP. When asking about the accuracy of

teacher effectiveness ratings based on the summative teacher ratings (Standards 1-5) and

student growth measures (Standard 6), specific questions in the category of "Attributes of

Evaluation Procedures" considered the six North Carolina Professional Teaching

Standards (Public Schools of North Carolina, 2013b). Other changes to survey items

asked for teacher experiences with content related to curriculum reforms and the

Common Core State Standards and for teacher perceptions of the impact of political

initiatives, such as RttT, on the evaluation process. Most survey items were Likert-type

scale items, although some questions asked for nominal data. Opportunities were

included for participants to add comments to selected questions on TEP (modified).

Providing commentary was optional. TEP (modified) is shown in full in Appendix F.

Stiggins and Duke (1988) drew two conclusions from the results of the original

TEP study, determining that the instrument provided quality data about teacher

perceptions of the evaluation experience and that a strong relationship existed between

the perceptions of teachers as to evaluation outcomes and the attributes of the evaluation

event. According to Stiggins and Duke, the goal of their teacher evaluation research was

to make the process of evaluation meaningful to participants, and they suggested that this

occurs for teachers when they perceive that the teacher evaluation process results in

personal and professional growth. The authors stipulated that TEP provides data about

teacher perceptions in these areas (Stiggins & Duke, 1988). Since the purpose of the

current study sought to determine teacher perceptions of the impact of NCTEP on

professional growth, teacher effectiveness, and attitudes toward teaching, a modified

version of the TEP survey is an appropriate tool for data collection.

Interview questions. Following their completion of the TEP (modified)

questionnaire, participants who provided contact information were asked to participate in

a follow-up interview. Individual interviews were conducted by the researcher to further investigate teacher perceptions of NCTEP. Interview questions were created by the researcher after quantitative and qualitative data analyses of survey responses were completed. When developing the interview questions, the researcher considered the "Survey Questions Used from the Interview Protocol (Teachers): Spring 2014" utilized by CERE-NC in the *Teacher and Principal Perceptions of the North Carolina Educator Evaluation System: Final Report* (Davis et al., 2015). Permission to use these questions from CERE-NC's interview protocol was granted by Cassandra Davis, lead author of the report. Appendix G documents Dr. Davis's consent. The CERE-NC interview protocol is located at http://cerenc.org/wp-content/uploads/2015/09/ES-TLEE-teacher-and-princ-perceptions-FINAL-9-15-15.pdf (Davis et al., 2015, pp. 32-35). A complete list of researcher developed, follow-up interview questions is located in Appendix H.

Procedures

The premise of this study was grounded in the theory that effective teacher evaluation holds the key to both teacher and school improvement (Stiggins & Duke, 1988). In the rapidly changing field of teacher evaluation (Hull, 2013), the ongoing study of teacher evaluation with particular attention to teacher perspectives is needed to gauge the capacity to which the evaluation system is meeting its primary purposes of improving teacher effectiveness and enhancing professional growth (NCDPI, 2015). To gain a richer understanding of teacher perceptions of NCTEP, a mixed-methods approach conducted in two phases was used to gather data. Data that were gathered detailed teacher experiences with NCTEP during their most recently completed annual evaluation cycle.

In the first phase of the study, TEP (modified) was administered to teachers in

Grades PreK-13 in a midsized school district in the foothills of North Carolina during

May 2016. Paricipation in the survey was voluntary. Principals within the district were

asked to introduce the opportunity for participation in the study to their faculty members

with the assurance that no schools would be linked to survey responses. All teachers in

the district received an email that described the study, gave directions for completing

TEP (modified), and contained the URL to the survey. The survey was administered

online through the Survey Monkey online service. To ensure that only certified teachers

completed the survey, two qualifying questions were asked at the beginning of the

survey:

1. I am a teacher, licensed by the Public Schools of North Carolina, and am

    subject to evaluation through the NCTEP; and

2. My most recent, completed annual evaluation cycle, which concluded in a

    summary evaluation conference, was for the: a. 2015-16 school year; b. 2014-

    2015 school year; c. Other (Give most recent evaluation year.); d. I have not

    completed an annual evaluation cycle.

If a participant answered "No" to question 1 or chose "d" for question 2, they were

thanked for their willingness to participate and exited from the survey. Respondents who

met the stipulated criterion continued with the TEP (modified) questionnaire. Survey

Monkey settings prohibited more than one completion of the survey from the same email

address, thus preventing multiple responses from the same participant.

The TEP (modified) survey was available for a 2-week period in May 2016. A

reminder was sent after the first week to teachers who had not responded to the survey.

A closing thank you was sent to participants at the end of the second week. Data were

compiled, and responses were downloaded in the IBM (Statistical Product and Service

Solutions) Statistics GradPack for Windows analysis.

TEP (modified) included Likert-type scale response items (Vagias, 2006), agreement rating scale items, and questions related to teacher experience levels in the classroom and their experiences with curriculum content and with NCTEP. The survey was based on Stiggins and Duke's (1988) study that administered the TEP Questionnaire in the fall of the 1986-1987 academic year; however, this researcher's study, including the questionnaire and analysis of data, replicated Stiggins and Duke's research.

Once respondents qualified as valid participants in the study, they completed the additional 57 items in TEP (modified) relating to their experiences with NCTEP. The final two questions allowed for the addition of any other information about the evaluation experience that the participant would like to specify in a narrative form and introduced an opportunity for respondents to continue participation in phase two of the study by providing a contact number or email. Narrative comments that were supplied for survey questions were analyzed. Information was categorized and reported as it related to the research questions. The researcher contacted teachers who supplied contact information in phase two of this study to schedule follow-up interviews.

Phase two of this study produced qualitative data from interviews of teachers who agreed to be included in further research about their evaluation experiences. As the TEP (modified) surveys were completed, the researcher compiled the list of phase two participants. In-depth interviews, which were unstructured and allowed for open responses by the interviewees (Frechtling & Sharp, 1997), were the format incorporated during phase two of the study. Interviews were held in July 2016 and were conducted either face-to-face or through phone interviews. Interviews were conducted on a one-to-one basis, and data were collected in voice recordings and transcribed verbatim. Verbal

permission was obtained from each participant prior to beginning recording. Transcripts of the interviews were offered to participants and were provided upon request.

The goal of interviewing during phase two was to ask probing, open-ended questions that captured the participant perceptions of their experiences with NCTEP in their own words (Frechtling & Sharp, 1997). Data in this phase of research consisted of information collected during the interview process. From transcriptions, data were reduced and categorized (Frechtling & Sharp, 1997) to focus on the topics of the three research questions. These questions were

1. What attributes of NCTEP impact teacher perceptions of professional growth?

2. How do the attributes of NCTEP impact teacher perceptions of teacher effectiveness?

3. How does NCTEP impact teacher attitudes toward teaching?

Conclusions about the data were carefully considered and reported in a narrative format (Frechtling & Sharp, 1997).

Delimitations

This study was limited to North Carolina licensed teachers who were working in the selected midsized school district of North Carolina. Participants were also limited to those teachers who were subject to evaluation through NCTEP and had completed an annual evaluation cycle which concluded in a summary evaluation conference. The survey was administered in May 2016 and did not include teachers who were in their first year of service or those who were working under a provisional license. Only teachers were surveyed in this study, so the findings may not be generalizable to others participating in North Carolina evaluation processes such as principals, mentors, assistant principals, instructional coaches, and central office personnel.

Factors impacting teacher evaluation that were not included in this study were the amount of resources, both human and fiscal, that are available for continual improvement of the evaluation system. Additionally, the constraints on observers' time, such as student discipline and building management, were not addressed as these factors could not be fully measured by the researcher. Finally, only teacher perceptions of the evaluation process were gathered through TEP (modified) and follow-up interviews. Perceptions of administrators, instructional coaches, central office personnel, or others connected to the teacher evaluation were not included in this research.

Limitations

In this study, the quantitative design utilized regression analysis of the four outcome rating questions in relation to the five key attributes of effective evaluation as determined by TEP (modified), which was based on Stiggins and Duke's (1988) earlier TEP questionnaire. While this quantitative methodology identified relationships between the outcome ratings and the attributes of the evaluation, the quantitative design allowed only for predictive relationships and did not identify casual relationships. The collection of qualitative data was necessary to support the findings generated from quantitative analysis and to answer the research questions.

This study was limited by the survey samplings. Only teachers who had completed NCTEP, including a summative evaluation, were considered in this research. Teachers who had not completed the full NCTEP were disqualified from participating in the online survey. Since participants for the follow-up interviews were identified from the survey, teachers who had not completed the full NCTEP were also excluded from this phase of the study. Survey results indicated that the majority of respondents to the survey received effective or highly effective status ratings. The perspectives of teachers in these

categories may differ from those receiving in need of improvement ratings. This factor may limit the study. Finally, this study was limited to one midsized school district in Western North Carolina, and this may limit the generalizability of the results.

As this study used a survey, validity of the results depended on the accuracy of teacher self-reporting. Teachers' potentially limited or biased remembrances of their evaluation experiences may have impacted the quality of the data collected. The time of year the survey occurred also may have limited this study as teachers were under higher stress near the end of the school year. The lack of inclusion of evaluator perspectives in relation to teacher self-reports limited the analysis of the accuracy of teacher self-reporting.

Summary

Notably, teacher quality has long been recognized as having a significant impact on student success (Marzano et al., 2011); however, determining the characteristics and practices of an effective teacher and how best to measure those components have proven to be both complex (Darling-Hammond et al., 1983) and controversial (Weisberg et al., 2009). While teachers generally agree that evaluation should have two primary functions, measuring teacher effectiveness and enhancing teacher professional growth (Marzano, 2012), the competing ideologies of stakeholders in the evaluation process make designing and implementing an evaluation system that leads to results in those arenas elusive (Taylor & Tyler, 2012).

This study attempted to determine teacher perceptions of NCTEP as it impacts professional growth and teacher effectiveness as well as identifying the impact of the overall evaluation experience on teacher attitudes toward teaching. Research data assisted in determining if and how NCTEP is accomplishing its intended purpose: "to

assess the teacher's performance in relation to the North Carolina Professional Teaching

Standards and to design a plan for professional growth" (Public Schools of North

Carolina, 2015, p. 4). Results of this study will help to inform ongoing implementation

planning and recommended staff development to ensure continuous improvement of

NCTEP. Additionally, this research identified attributes of NCTEP that teachers perceive

as impacting professional growth and specified if and how NCTEP attributes impact

teacher effectiveness and attitudes toward teaching. Study results indicated attributes

need to be considered for revision so teachers participate in meaningful evaluation

experiences they perceive promote personal and professional growth.

Chapter 4: Results

Introduction

The purpose of this study was to examine the impact of NCTEP on teacher

perceptions of professional growth, teacher effectiveness, and attitudes toward teaching.

Marzano (2012) stated that teachers generally agree that evaluation should have two

primary functions: measuring teacher effectiveness and enhancing teacher professional

growth; however, whether the primary goal of evaluation should be growth or

accountability has been influenced by political and public opinion (Darling-Hammond,

2014). Over the past several decades, teacher evaluation reform efforts have evolved

from focusing on teacher growth (Stiggins & Duke, 1988) to an increased emphasis on

teacher evaluation for accountability (Darling-Hammond, 2014). The purposes of

NCTEP reflect both accountability and growth goals (Public Schools of North Carolina,

2015). Data reported in this chapter provide insight into teacher perceptions of how

NCTEP is accomplishing these growth and accountability objectives as well as how the

evaluation process impacts their attitudes toward teaching.

Participants in this study were certified teachers who were employed during 2015-

2016 by a midsized school district in Western North Carolina. Teachers were categorized

by the district as elementary (K-8), high school (9-above), or other. Counselors, school

psychologists, and media coordinators were considered certified teachers and were

included in the other category for reporting purposes. The total number of 817 certified

teachers employed by the district was composed of 397 elementary teachers, 198 high

school teachers, and 222 teachers in the other category (B. Johnson, personal

conversation, April 12, 2016).

Data were collected in two phases of research. Both quantitative and qualitative

data were generated in the study through online survey responses to the TEP (modified)

questionnaire and from one-on-one interviews. Quantitatively, survey data were

analyzed to determine correlations between the five attributes of effective evaluation and

outcome ratings of the overall quality of the evaluation process and its impact on

professional growth, teacher effectiveness, and attitudes toward teaching. Correlations

were also calculated for data from respondents who reported receiving effective and

highly effective status ratings. The low number of respondents reporting an in need of

improvement status eliminated this category from analysis. Frequency data were

generated and analyzed for all survey responses. Qualitative data were retrieved from

additional comments given in the online survey and from follow-up interviews.

The following research questions guided this study:

1. What attributes of NCTEP impact teacher perceptions of professional growth?

2. How do the attributes of NCTEP impact teacher perceptions of teacher
   effectiveness?

3. How does NCTEP impact teacher attitudes toward teaching?

Data are reported and analyzed in the context of these research questions and will

build on an organizational framework that considers the five key attributes of effective

evaluation (teacher, evaluator, evaluation process, feedback, and context of the

evaluation) as determined by Stiggins and Duke (1988).

Data Analysis

Quantitative data. Quantitative data analyzed in this study were collected

through TEP (modified), an online survey, distributed to all certified teachers in the

district. An average response rate of 22.5% was received, with response rates declining

slightly throughout the survey from 192 responses on question 21 to 179 responses on

question 61. Response rates are reported on frequency charts located in Appendix I.
Participants were qualified for the survey by their responses to "I am a teacher, licensed
by the Public Schools of North Carolina and subject to an annual evaluation in the
NCTEP." Question 3 affirmed that the participant had completed an annual evaluation
cycle. Respondents who did not meet the stipulated criterion were exited from the online
survey.

Nominal data. Nominal level data were gathered through NCTEP (modified)
that provided descriptive information about respondents. Consideration of teacher status
ratings; teaching experience (years); and grade level taught at time of most recent,
completed evaluation cycle were included in data analysis. Although other nominal level
data were collected, they were excluded from data analysis due to discrepancies in
responses identified by the researcher or difficulty in establishing a quantifiable
relationship to the outcome ratings questions. Categories of data excluded from
quantitative analysis include teacher experience with the North Carolina Standard Course
of Study and with the TPAI-R evaluation instrument and details about observation types,
observers, and PDP providers.

Ordinal data. NCTEP (modified) provided information about teacher
perceptions of the key attributes of effective evaluation (teacher, evaluator, evaluation
process, feedback, and context of the evaluation) as identified by Stiggins and Duke
(1988) in relation to outcome ratings. Attributes were examined through teacher
responses to 43 Likert scale-type items that ranked descriptors within each attribute
category. Ordinal level data from NCTEP (modified) were used to calculate the mean
response for each of the five categories of attributes of effective evaluation. The mean
response for each attribute was correlated to the outcome ratings generated in questions 4

through 7, which reported teacher perceptions of the overall quality of the evaluation and

impact of the evaluation process on teacher growth (changes in teaching

behaviors/strategies), teacher effectiveness (changes in understanding of the

teaching/learning process), and attitudes toward teaching. Outcome ratings as conveyed

through teacher responses were criterion variables for the quantitative phase of this study.

Predictor variables were the mean responses calculated for each attribute.

Pearson correlation coefficents by overall response. Pearson product-moment

correlation coefficients were used to examine relationships among the five key attributes

of effective evaluation as identified by Stiggins and Duke (1988) and the four outcome

ratings. Tables 4 and 5 illustrate the five key attributes in relation to outcome ratings for

all responses. Table 4 shows the correlation results for all survey responses between

outcome ratings and attributes of teacher, evaluator, evaluation procedures.

Table 4

Pearson r Correlations Between Outcome Ratings and Attributes of Teacher, Evaluator, and Evaluation Procedures – All Responses

Outcome Ratings	Teacher Attributes			Evaluator Attributes			Evaluation Procedure Attributes		
	r	*p*	N	*R*	*P*	N	*r*	*p*	N
Overall Quality	.140	.053	192	.204**	.005	189	.308**	.000	182
Impact on Teaching Behaviors/Strategies (Growth)	.031	.672	189	.161*	.028	186	.342**	.000	179
Impact on Understanding of Teaching/Learning (Effectiveness)	.041	.577	191	.200**	.006	188	.402**	.000	181
Impact on Attitudes toward Teaching	.073	.315	190	.123	.094	187	.296**	.000	180

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

Table 5 shows the correlation results for all survey responses between outcome ratings and attributes of feedback and context of the evaluation.

Table 5

Pearson r Correlations Between Outcome Ratings and Attributes of Feedback and Context of the Evaluation – All Responses

Outcome Ratings	Feedback Attributes			Context Attributes		
	R	*p*	N	*r*	*p*	N
Overall Quality	.308**	.000	182	.349**	.000	182
Impact on Teaching Behaviors/Strategies (Growth)	.342**	.000	179	.468**	.000	179
Impact on Understanding of Teaching/Learning (Effectiveness)	.402**	.000	181	.471**	.000	181
Impact on Attitudes toward Teaching	.296**	.000	180	.371**	.000	180

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

Data show that a weak to moderate, positive correlation exists between the overall quality of the evaluation process and the attributes of evaluators ($r(187)=.204$, $p<.005$); evaluation procedures ($r(180)=.308$, $p<.001$); feedback ($r(180)=.308$, $p<.001$); and context of the evaluation ($r(180)=.349$, $p<.001$). Correlations indicate a significant linear relationship between the overall quality of the evaluation process and these attributes.

Outcome ratings for measures related to teacher growth (changes to teacher behaviors and strategies) showed a weak to moderate, positive correlation to attributes of evaluators ($r(184)=.161$, $p<.028$); evaluation procedures ($r(177)=.342$, $p<.001$); feedback ($r(177)=.342$, $p<.001$); and context of the evaluation ($r(177)=.468$, $p<.001$). A significant linear relationship exists between teacher growth measures and these four attributes.

Outcome measures related to teacher effectiveness (the impact of the teacher evaluation process on understanding of the teaching/learning process) were found to have a weak to moderate, positive correlation to the attributes of evaluators ($r(186)=.200$, $p<.006$); evaluation procedures ($r(179)=.402$, $p<.001$); feedback ($r(179)=.402$, $p<.001$); and context of the evaluation ($r(179)=.471$, $p<.001$). Correlations indicated a significant linear relationship between outcome ratings for teacher effectiveness and four of the five attribute categories.

A moderate, positive correlation was found between attitudes toward teaching and the attributes of evaluation procedures ($r(178)=.296$, $p<.001$); feedback ($r(178)=.296$, $p<.001$); and context of the evaluation ($r(178)=.371$, $p<.001$). A significant linear relationship exists between attitudes toward teaching and three of the five attributes of effective evaluation.

A Pearson correlation was calculated that examined the relationship between the four outcome ratings and teacher attributes. A weak correlation that was not significant was found for teacher attributes and outcome ratings of overall quality ($r(190)=.140$, $p>.05$); impact on teaching behaviors/strategies ($r(187)=.672$, $p>.05$); impact on understanding of teaching/learning process ($r(189)=.577$, $p>.05$); and impact on attitudes toward teaching ($r(188)=.315$, $p>.05$). Teacher attributes are not related to the outcome ratings as reported by the TEP (modified) survey.

Pearson correlation coefficients by status ratings. A component of NCTEP includes the assignment of status ratings (in need of improvement, effective, and highly effective) to teachers (Public Schools of North Carolina, 2015). Data were sorted by teachers who were rated effective and highly effective. Analysis of data by teachers who were rated in need of improvement were not considered as only two participants

identified themselves in this category. Pearson product-moment correlation coefficients

were used to examine relationships between outcome ratings and attributes of evaluation

for effective and highly effective teachers. Tables 6 and 7 illustrate these relationships.

Table 6 shows the correlation results between outcome ratings and attributes of teacher,

evaluator, and evaluation procedures.

Table 6

Pearson r Correlations Between Outcome Ratings and Attributes of Teacher, Evaluator and Evaluation Procedures – Effective Teacher Responses

Effective Teacher Outcome Ratings	Teacher Attributes			Evaluator Attributes			Evaluation Procedure Attributes		
	r	*P*	N	*r*	*P*	N	*r*	*P*	N
Overall Quality	.159	.209	64	.150	.250	61	.324*	.013	58
Impact on Teaching Behaviors/Strategies (Growth)	.074	.563	63	.202	.122	60	.499**	.000	57
Impact on Understanding of Teaching/Learning (Effectiveness)	.086	.505	63	.158	.228	60	.470**	.000	57
Impact on Attitudes toward Teaching	.282*	.025	63	.238	.067	60	.430**	.001	57

**Correlation is significant at the 0.01 level (2-tailed).
 *Correlation is significant at the 0.05 level (2-tailed).

Table 7 shows the correlation results between outcome ratings and attributes of

feedback and context of the evaluation for effective teachers.

Table 7

Pearson r Correlations Between Outcome Ratings and Attributes of Feedback and Context of the Evaluation – Effective Teacher Responses

	Feedback Attributes			Context Attributes		
Effective Teacher Outcome Ratings	*R*	*p*	N	*r*	*p*	N
Overall Quality	.324*	.013	58	.454**	.000	58
Impact on Teaching Behaviors/Strategies (Growth)	.499**	.000	57	.682**	.000	57
Impact on Understanding of Teaching/Learning (Effectiveness)	.470**	.000	57	.648**	.000	57
Impact on Attitudes toward Teaching	.430**	.001	157	.535**	.000	57

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

As denoted in Tables 6 and 7, responses from teachers who received effective ratings showed weak, positive correlations between the outcome rating, impact on attitudes toward teaching, and teacher attributes ($r(61)=.025$, $p<.01$). Moderate, positive correlations were demonstrated for the four outcome ratings (overall quality, impact on teaching behaviors/strategies, impact on understanding of the teaching/learning process, and impact on attitudes toward teaching) in relationship to attributes of evaluation procedures, feedback, and context of the evaluation. While still in the moderate range, correlations between outcome ratings and context of the evaluation attributes were stronger than relationships between the four outcome ratings and the attributes of teacher, evaluator, evaluation procedures, and feedback.

Table 8 shows correlations of outcome ratings and attributes of teacher, evaluator and evaluation procedures by highly effective teachers.

Table 8

Pearson r Correlations Between Outcome Ratings and Attributes of Teacher, Evaluator, and Evaluation Procedures – Highly Effective Teacher Responses

Highly Effective Teacher Outcome Ratings	Teacher Attributes			Evaluator Attributes			Evaluation Procedure Attributes		
	r	p	N	R	P	N	r	p	N
Overall Quality	.146	.123	113	.226*	.016	113	.270**	.004	110
Impact on Teaching Behaviors/Strategies (Growth)	.047	.626	112	.219	.020	112	.261**	.006	109
Impact on Understanding of Teaching/Learning (Effectiveness)	.078	.413	113	.281**	.003	113	.352**	.000	110
Impact on Attitudes toward Teaching	.021	.825	112	.205*	.030	112	.300**	.002	109

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

Table 9 shows correlations of outcome ratings and attributes of feedback and

context of the evaluation by highly effective teachers.

Table 9

Pearson r Correlations Between Outcome Ratings and Attributes of Feedback and Context of the Evaluation – Highly Effective Teacher Responses

Highly Effective Teacher Outcome Ratings	Feedback Attributes			Context Attributes		
	R	P	N	r	p	N
Overall Quality	.270*	.004	110	.270*	.004	110
Impact on Teaching Behaviors/Strategies (Growth)	.261**	.006	109	.310**	.001	109
Impact on Understanding of Teaching/Learning (Effectiveness)	.352**	.000	110	.338**	.000	110
Impact on Attitudes toward Teaching	.300**	.002	109	.318**	.001	109

**Correlation is significant at the 0.01 level (2-tailed).
 *Correlation is significant at the 0.05 level (2-tailed).

As illustrated by Tables 8 and 9, responses of teachers who were rated highly effective showed no significant relationship between the attributes of teachers and any of the four outcome ratings (overall quality, impact on teaching behaviors/strategies, impact on understanding of the teaching/learning process, and impact on attitudes toward teaching). All other attributes (evaluators, evaluation procedures, feedback, and context of the evaluation) demonstrate moderate, positive relationships to the all four outcome ratings.

Frequency distributions and crosstabulations. Frequency distribution charts were created for questions 4 through 60 from the TEP (modified) questionnaire. A complete list of frequency charts is included in Appendix I. Close analysis was given to frequency charts for specific questions, and data were used to develop follow-up interview questions and to gain insight into teacher perceptions about the evaluation

process. Of note, questions 8 and 10-15, which related to teacher attributes, were explored in detail as this category did not show relationships to outcome ratings using Pearson correlation coefficients. Analysis of ratings revealed that an overall average of 78.7% of teachers chose a 4 or 5 on the Likert scale as their response to teacher attribute questions. Question 53 examined the use of NCEES for enhancing the evaluation process and allowed for optional comments. Frequency data showed that approximately 32% of responses gave low rankings (1-2) to NCEES as enhancing the evaluation process; approximately 33% gave mid rankings (3); and approximately 35% gave high rankings (4-5). Low rankings were described by not at all; and high rankings were considered to enhance the process a great deal. Responses were normally distributed for this question.

Frequency data were also analyzed for question 58, which asked teachers to identify the intended role of evaluation. Results from the survey showed that slightly more than one third (34.3%) reported a balance between accountability and growth, choosing a 3 on the 5-point Likert scale. Less than one fourth (23.7%) of the respondents noted that the role leaned more toward accountability, while the majority of responses (42%) leaned more toward teacher growth in their responses.

Crosstabulations of survey items were calculated to compare teacher rankings of outcome ratings to teacher experience levels and grade-level assignments at the time of their most recent, completed evaluation cycle. Crosstabulation data were also created to further examine respondent comments on specified survey items. To simplify reporting, the rankings for outcome ratings were condensed from a 10-point scale to three descriptive categories: low ranking (0-3), mid ranking (4-5), and high ranking (6-9). Teacher experience levels were condensed and categorized by 10 years or less, 11-19 years, and 20 years or more. Additionally, grade-level assignments were condensed and

classified as PreK-5, 6-8, 9 and above, and other. Data in which teachers specified their

grade-level assignments were not easily classified and, therefore, were included in the

other category. Likert scale-type items were not modified for reporting purposes and

utilized a five-point scale with words or phrases to describe opposite ends of the rating

scales.

Table 10 shows rankings for the evaluation process's impact on teaching

strategies and behaviors by teachers' years of experience.

Table 10

Outcome Rating (Teaching Strategies/Behaviors) Rankings by Teachers' Experience (Years)

	All Responses By Experience		Low Ranking (0-3)		Mid Ranking (4-5)		High Ranking (6-9)	
	N	%	N	%	N	%	N	%
10 Years or less	60	31.7	15	25.0	21	35.0	24	40.0
11-19 Years	67	35.5	18	26.8	29	43.3	20	29.9
20 Years or more	62	32.8	24	38.7	21	33.9	17	27.4
Total	189	100.0	57	30.2	71	37.6	61	32.2

Approximately 39% of teachers with 20 years or more of experience indicated a

low ranking impact of the evaluation process on changes to their teaching strategies/

behaviors, whereas 40% of teachers with 10 years or less gave a high ranking for impact

on teaching strategies/behaviors. Data show that overall teachers with less experience

report lower percentages in the low ranking category by up to approximately 13%.

Teachers with more experience report lower percentages in the high ranking category by

an approximate 13% difference from teachers with the most experience. Teachers with

11-19 years of experience report the highest percentage (43.3%) in the mid ranking category.

Table 11 illustrates rankings for the evaluation process's impact on teaching the understandings of the teaching/learning process by teachers' years of experience.

Table 11

Outcome Rating (Understanding of Teaching/Learning Process) Rankings by Teachers' Experience (Years)

	All Responses By Experience		Low Ranking (0-3)		Mid Ranking (4-5)		High Ranking (6-9)	
	N	%	N	%	N	%	N	%
10 Years or less	61	31.9	20	32.8	18	29.5	23	37.7
11-19 Years	68	35.6	19	27.9	30	44.1	19	28.0
20 Years or more	62	32.5	32	51.6	11	17.8	19	30.6
Total	191	100.0	71	37.2	59	30.9	61	31.9

Analysis of crosstabs of the impact of teacher evaluation processes on the understanding of the teaching and learning process and respondents' years of teaching experience show that teachers with 20 years or more experience have the highest percentage of low rankings (51.6 %) for this outcome rating. This is 18.8% higher than teachers in the 10 years or less group. Teachers with 11-19 years of experience have the highest percentage (44.1%) of mid ranking responses, which is 14.6% higher than the 10 years or less group and 26.3% higher than the 20 years or more group. Teachers with 10 years or less of experience most often (37.7%) give a high ranking to the evaluation's impact on their understanding of the teaching/learning process.

Table 12 shows crosstabs for the impact of the evaluation process on teacher attitudes toward teaching based on their years of teaching experience.

Table 12

Outcome Rating (Attitude) Rankings by Teachers' Experience (Years)

	All Responses By Experience		Low Ranking (0-3)		Mid Ranking (4-5)		High Ranking (6-9)	
	N	%	N	%	N	%	N	%
10 Years or less	61	32.1	12	19.7	21	34.4	28	45.9
11-19 Years	67	35.3	23	34.3	21	31.4	23	34.3
20 Years or more	62	32.6	25	40.3	14	22.6	23	37.1
Total	190	100.0	60	31.6	56	29.5	74	38.9

Table 12 shows crosstabs for the impact of the evaluation process on teacher attitudes toward teaching based on their years of teaching experience. Data show that teachers with the most experience were the most likely (40.3%) to report a low impact on their attitudes toward teaching, whereas teachers with the least experience were most likely (45.9%) to report a high impact. A smaller difference in percentages between these two groups is evident in high ranking responses (8.8% difference) than in low ranking responses (20.6% difference). Approximately one third (34.4%) of teachers with 10 years or less and one third (31.4%) of teachers with 11-19 years of experience gave the evaluation process a mid-level ranking as to its impact on their attitudes toward teaching. Crosstabulations show that teachers with 10 years or less experience are the least likely (19.7%) to report a low ranking impact of the evaluation process on their attitudes toward teaching.

Table 13 shows outcome ratings for impact on teaching strategies/behaviors according to grade-level assignment during the last completed evaluation cycle.

Table 13

Outcome Rating (Teaching Strategies/Behaviors) Rankings by Grade Level Assignment

	All Responses By Experience		Low Ranking (0-3)		Mid Ranking (4-5)		High Ranking (6-9)	
	N	%	N	%	N	%	N	%
PreK-5	52	27.7	12	23.0	20	38.5	20	38.5
6-8	51	27.1	12	23.5	21	41.2	18	35.3
9 and above	55	29.2	26	47.3	18	32.7	11	20.0
Other	30	16.0	6	20.0	12	40.0	12	40.0
Total	188	100.0	56	29.8	71	37.8	61	32.4

Teachers in Grades 9 and above report the highest percentage (47.3%) of low rankings for the impact of teaching strategies/behaviors, demonstrating a 24.3% difference from PreK-5 teachers; 23.8% difference from 6-8 teachers; and a 27.3% difference to other teachers. Forty percent of teachers in the other category reported high rankings of the evaluation's impact on changes to teaching strategies/behaviors. This percentage was 1.5% higher than Grades 6-8 teachers and 20% higher than teachers of Grades 9 and above. Mid ranking data demonstrated little variance in percentages between grade-level groups, with less than 3-9% differences between the four grade-level assignments.

Table 14 illustrates rankings for the evaluation process's impact on the understanding of the teaching/learning process by grade-level assignment during the last completed evaluation cycle.

Table 14

Outcome Rating (Understanding of Teaching/Learning Process) Rankings by Grade Level Assignment

	All Responses By Experience		Low Ranking (0-3)		Mid Ranking (4-5)		High Ranking (6-9)	
	N	%	N	%	N	%	N	%
PreK-5	53	27.9	13	24.5	21	39.6	19	35.9
6-8	51	26.8	19	37.3	14	27.4	18	35.3
9 and above	56	29.5	30	53.6	12	21.4	14	25.0
Other	30	15.8	8	26.7	12	40.0	10	33.3
Total	190	100.0	70	36.8	59	31.1	61	32.1

Data show that 53% teachers in Grades 9 and above report a low impact ranking of the evaluation process on their understanding of the thinking/learning process, whereas 25% of this same group reports a high ranking impact. Data indicate that a similar percentage of PreK-5 teachers (35.9%) and 6-8 teachers (35.3%) report a high ranking impact on the understanding of the teaching and learning process. Forty percent of teachers in the other category gave a mid ranking to this outcome rating question.

Table 15 shows crosstabulations for the impact of the evaluation process on teacher attitudes toward teaching based on their grade-level assignment during the last completed evaluation cycle.

Table 15

Outcome Rating (Attitude) Rankings by Grade Level Assignment

	All Responses By Experience		Low Ranking (0-3)		Mid Ranking (4-5)		High Ranking (6-9)	
	N	%	N	%	N	%	N	%
PreK-5	54	28.6	16	29.6	17	31.5	21	38.9
6-8	49	25.9	12	24.5	13	26.5	24	49.0
9 and above	56	29.6	25	44.6	14	25.0	17	30.4
Other	30	15.9	6	20.0	12	40.0	12	40.0
Total	189	100.0	59	31.2	56	29.6	74	39.2

A higher percentage of teachers (44.6%) in Grades 9 and above reports low rankings for the impact of the evaluation process on their attitudes toward teaching, which is approximately 15% to 25% higher than any other group. Forty percent of teachers in the other category gave a mid ranking to the evaluation's impact on their attitudes. This is a higher percentage by approximately 8% to 15% than that reported by any other grade-level assignments. Forty-nine percent of Grades 6-8 teachers reported a high ranking as to the impact of the evaluation process on their attitudes toward teaching.

Qualitative data. Qualitative data for this study were collected through two methods. First, respondents were given the opportunity to make optional comments about survey questions 4, 5, 6, 7, 53, and 61 as part of the NCTEP (modified) online survey. Their comments were collected anonymously. The percentage of participants who provided additional narratives to the survey questions varied from 8.2% to 21.2%. Question 53, concerning NCEES, generated the fewest responses: 15 of 182. Question 61, which asked for other dimensions to consider in the evaluation process, generated the greatest percentage of comments (21.2%) with 38 comments of 179 responses.

A second method of qualitative data collection was generated from a final question on NCTEP (modified) that asked participants to voluntarily participate in follow-up interviews with the researcher. Twenty-four respondents provided contact information and expressed a willingness to participate. Face-to-face or phone interviews were scheduled with 14 of the 24 volunteers, and interviews were recorded verbatim and transcribed by a transcriptionist for analysis. Respondents were offered a copy of the transcript.

Comments from NCTEP (modified). Comments from NCTEP (modified) were printed and color coded to identify themes. Themes were then categorized by the five attributes of effective evaluation (teacher, evaluator, evaluation procedure, feedback, and context of the evaluation) or were identified as not connected. Comments were also categorized by their tone of the response (positive, negative, both positive and negative, neither positive nor negative, or unclear). Single comments may have included multiple themes; therefore, the number of themes may be more than the number of individual comments. Likewise, a theme could be related to more than one attribute.

Frequency codes were created by the researcher to determine the strength of responses. The frequency of themes and their connections to attributes were used to identify strength. A strong (S) theme was mentioned 11 or more times; moderate (M) themes were mentioned six to 10 times; weak (W) themes were mentioned one to five times. Themes were considered no relation (NR) if they were not mentioned. Strength codes for themes and for attribute totals are identified by their abbreviations on the following tables.

Table 16 details the tone of comments for question 4, which asked teachers to rate the overall quality of the evaluation process.

Table 16

Tone of Responses for Overall Quality of NCTEP

	Positive (+)	Negative (-)	Both +/-	Neither +/-	Unclear
Tone of Comments	1-W	30-S	6-M	1-W	1-W

Thirty-nine comments concerning the quality of the evaluation were entered. Comments were overwhelming negative at 77%; although some respondents saw potential in the process, even with reservations. One teacher stated, "The design is well planned, but implementation is lacking." The negativity of the comments is in contrast to the quantitative data that showed almost half (47.2%) of the 197 responses to the outcome rating leaned toward the descriptor high quality; and 34.5% of the 197 responses were neutral, leaving less than 20% of the total group choosing a low rating for quality of the evaluation.

Table 17 presents the themes identified in teacher comments about the overall quality of the evaluation process.

Table 17

Themes and Connections to Attribute Categories for the Overall Quality of NCTEP

Themes: Elements that contribute to Quality	# of Times Mentioned	Connected to Attribute (Teacher, Evaluator, Procedures, Feedback, Context, Not Connected)					
		T	E	P	F	C	NC
Biased/unfair	15-S		10	5			
Administrator/evaluator role	12-S		12				
Inability to reflect all aspects of teaching	10-M	1		6		3	
Time Issues	7-M			6		1	
State oversight	3-W					3	
NCEES	3-W					3	
Standard 6	3-W			3			
Teacher reflection	1-W	1					
Totals		2-W	22-S	20-S	0-NR	10-M	0-NR

Two themes emerged as strong in comments about the elements of the evaluation that contribute to its quality: biased/unfair and the administrator/evaluator role. The attributes with the strongest ratings were evaluator and procedures. One comment exemplified the impact of these attributes by stating, "The evaluation process various (sic.) enormously based on who is selected to be your evaluator. If you get someone who is communicative and helpful, the system works fine. But if you get a taciturn or cryptic evaluator, good luck."

Table 18 illustrates the tone of responses for question 6, in which teachers were asked to rate the impact of the evaluation process on changes to their teaching strategies and behaviors; i.e., teacher growth.

Table 18

Tone of Responses for Impact of NCTEP on Teacher Growth

	Positive (+)	Negative (-)	Both +/-	Neither +/-	Unclear
Tone of Comments	4-W	14-S	4-W	4-W	0-NR

More than 50% of comments about the impact of the evaluation process on teacher growth were negative. Teachers remarked that "the process is too subjective," "fails to measure basic things, like working well in a team, being reliable, punctual," and "is difficult to move through in that there are too many things to click on."

Table 19 reports themes that were found in comments about the impact on changes to teaching strategies and behaviors.

Table 19

Themes and Connections to Attribute Categories for the Impact on Changes to Teaching Strategies/Behaviors

Themes – Elements that contribute to growth	# of Times Mentioned	Connected to Attribute (Teacher, Evaluator, Procedures, Feedback, Context, Not Connected)					
		T	E	P	F	C	NC
Professional growth unrelated to/not impacted by NCTEP	17-S						17
NCTEP – Feedback	8-M				8		
NCTEP – Procedures	6-M			6			
NCTEP – Teacher	2-W	2					
Totals		2-W		6-M	8-M		17-S

The theme designating that professional growth is unrelated to/not impacted by NCTEP was strong, as was the not connected attribute category. The attributes of procedures and feedback were moderately strong. One teacher reported that "As a veteran teacher, I care about evaluations but feel my overall impact on education is overlooked." Another stated, "I don't feel that I get critical feedback that would help me change my instructional behavior."

Table 20 reports the tone of comments for question 7. This question asked respondents to rate the impact of the evaluation process on their understanding of the teaching and learning process; i.e., measures of effectiveness.

Table 20

Tone of Responses for Impact of NCTEP on Effectiveness

	Positive (+)	Negative (-)	Both +/-	Neither +/-	Unclear
Tone of Comments	2-W	12-S	2-W	3-W	1-W

Sixty percent of comments about the impact of the evaluation on effectiveness were negative. As one teacher commented, "The evaluation process has nothing to do with my understanding of the evaluation process."

Table 21 details themes that were identified as impacting teacher understanding of the teaching and learning process.

Table 21

Themes and Connections to Attribute Categories for the Impact of NCTEP on Effectiveness

Themes – Elements impacting effectiveness	# of Times Mentioned	Connected to Attribute (Teacher, Evaluator, Procedures, Feedback, Context, Not Connected)					
		T	E	P	F	C	NC
Little or no impact due to NCTEP	18-S		1	4	2	2	9
NCTEP – Procedures	14-S			14			
NCTEP – Evaluator	4-W		4				
NCTEP – Feedback	4-W				4		
NCTEP – Context	4-W					4	
NCTEP – Teacher	1-W	1					
Totals		1-W	5-W	18-S	6-M	6-M	9-M

Comments demonstrated two strong themes concerning the impact of the evaluation process: little or no impact due to NCTEP and procedures. A strong connection to the evaluation procedures attribute was also indicated. Comments supported these themes, stating, "the evaluation has next ot (sic.) no impact on my instructional decisions" and "it is too general and some things like 'global awareness' are hard to prove in a hit and miss observation."

Table 22 reports the tone of reponses for question 5, which asked teachers to rate the impact that the teacher evaluation process had on their attitudes toward teaching.

Table 22

Tone of Responses for the Impact of NCTEP on Attitudes Toward Teaching

	Positive (+)	Negative (-)	Both +/-	Neither +/-	Unclear
Tone of Comments	3-W	20-S	2-W	4-W	1-W

Thirty comments were entered concerning attitudes toward teaching. Sixty-seven percent of these comments were negative. One teacher stated, "I do not believe the process in and of itself generates any impact. Again, it is the professionalism and effectiveness of the evaluator, not the tool, that determines its impact." Another remarked, "I am at the end of my career and if I were a new teacher I would seriously consider leaving the profession."

Table 23 details the themes found in teacher comments about the impact that the evaluation process had on their attitudes toward teaching.

Table 23

Themes and Connections to Attribute Categories for the Impact of NCTEP on Attitudes Toward Teaching

Themes – Elements that impact attitudes	# of Times Mentioned	Connected to Attribute (Teacher, Evaluator, Procedures, Feedback, Context, Not Connected)					
		T	E	P	F	C	NC
Has no impact/little connection to teaching	13-S						13
Increased stress	9-M		7	2			
Busy work/one more thing to do	8-M					8	
Influenced by administration	6-M		6				
Encourages reflection/ communication	2-W	1		1			
Positive attitudes about self	2-W	1	1				
Totals		2-W	14-S	5-W	0-NR	8-M	13-S

The theme, has no impact/little connection to teaching, received a strong rating. Attributes of the evaluator also received a strong code but was only one point above the not connected category. The attribute of evaluator was, however, reflected in this teacher's statement: "made me want to quit teaching and left me with no confidence with myself even though the previous year's evaluation was great, this year's was a total bomb even by the same observer." Themes related to reflection/communication and positive attitudes about self were weak.

Table 24 reports the tone of comments for question 53, concerning the role NCEES has in enhancing the evaluation process. The tone of most comments was negative (60%). In contrast, responses to the Likert scale item was normally distributed, with approximately 32% of answers demonstrating that teachers did not perceive that NCEES enhanced NCTEP; approximately 35% reporting that NCEES did enhance NCTEP; and 33% choicing the mid-point, a 3, to describe how NCEES enhanced NCTEP.

Table 24

Tone of Responses for Role of NCEES in Enhancing NCTEP

	Positive (+)	Negative (-)	Both +/-	Neither +/-	Unclear
Tone of comments	2-W	9-M	1-W	2-W	1-W

Table 25 examines teacher comments about NCEES and how they are connected to the attributes of effective teaching.

Table 25

Themes and Connections to Attribute Categories for the Role of NCEES in Enhancing NCTEP

Themes – Use of NCEES to enhance NCTEP	# of Times Mentioned	Connected to Attribute (Teacher, Evaluator, Procedures, Feedback, Context, Not Connected)					
		T	E	P	F	C	NC
Did not enhance process/ waste of time/confusing	12-S			4	2	1	5
Enhanced process	3-W		1	1		1	
Total		0-NR	1-W	5-W	2-W	1-W	5-W

A strong frequency of responses was found for the theme that stated the use of NCEES did not enhance the process, was confusing, or was a waste of time. The frequencies for connections to all attributes was weak or not mentioned. Three comments clarified teacher perceptions of the usefulness of NCEES by reporting: "The online platform is not user friendly"; "There is not really anyway to voice your opinion about your observation without jeopardizing your current position"; and "very confusing."

Question 61 asked, "Are there other dimensions of you as a teacher, the nature of the performance data collected, the nature of the feedback, the evaluation context, or other factors that you think are related to the success (or lack of success) of your past teacher evaluation experiences that should be included in this study?" This question generated 38 responses and reflected respondent thoughts about the current attributes in addition to other factors to consider in relation to NCTEP. Table 26 lists the frequency of themes reported by participants with their strength code.

Table 26

Other Dimensions to Include in Further Study of NCTEP

Themes – Other dimensions to include in study of NCTEP	# of Times Mentioned
Student factors	22-S
Teacher factors, including relational qualities	21-S
Evaluation Procedures	8-M
Family/Community Characteristics	5-W
School climate	4-W
Quality of PD	3-W
Quality of the Evaluator	2-W

Teachers commented that the "extra responsibilities of EC teaching should be considered," as well as "the prior condition of the environment." Other teachers pointed out that "extra school activities, relationships with students," "student absences and apathy" should be included in the study of teacher evaluation and noted that "continuous support and professional development is lacking."

Follow-up interviews. After data analysis of TEP (modified) was completed, interview questions were developed, and follow-up interviews with teachers were scheduled and conducted. Correlation data, frequency distributions, and narrative responses to specific survey items were considered in the development of follow-up interview questions. Teachers were asked to give their perspectives on the impact of NCTEP on their professional growth, teacher effectiveness, and their attitudes toward teaching. Interview questions are located in Appendix H. Fourteen teachers were interviewed. All interviewees responded to the same questions. If deemed appropriate in the scope of the study, an interviewee was asked to provide more details about his or her unique experience with the NCTEP process.

Teaching experience of the interviewees ranged from 6 to 27 years. Five teachers reported 0-10 years of teaching experience; four had 11-19 years of experience; and five teachers had 20+ years of teaching experience. Three of the teachers were assigned to Grades PreK-5; three worked in Grades 6-8; seven worked in Grades 9 and above; and one was considered in the other category as described previously.

When asked about their experiences with NCTEP, all 14 teachers reported being observed; five reported having served as a peer observer; three discussed their past experiences with TPAI in comparison to NCTEP; and one had trained and will serve as an evaluator for the ASW component of NCTEP. One teacher stated that she had worked with NCDPI in 2005 to develop the rubric that is currently used in NCTEP.

As this study is concerned with identifying the attributes of NCTEP that teachers perceive as contributing to their professional growth, interviewee comments were coded and categorized for themes; and then the themes were categorized by the five attributes of effective evaluation. Teachers were asked to identify the attributes that were the most and least likely to improve professional practice. Strength codes were created based on the frequencies of responses. A strong (S) theme was mentioned 11 or more times; moderate (M) themes were mentioned six to 10 times; weak (W) themes were mentioned one to five times. Themes were considered no relation (NR) if they were not mentioned. Strength codes for themes and for attribute totals are identified by their abbreviations on the following tables.

Table 27

Themes and Attribute Categories that are Most Likely to Impact on Professional Growth

Themes – Elements of NCTEP most likely to impact growth	# of Times Mentioned	Connected to Attribute (Teacher, Evaluator, Procedures, Feedback, Context, Not Connected)					
		T	E	P	F	C	NC
Quality of feedback/ productive conversations	18-S		3		15		
Promotes self-reflection/goal-setting	12-S	13		1	2		
Gives focus through standards/expectations	10-M	1		15	1	1	
Confidence in the evaluator	10-M		10				
Gives focus through PDP/ PD	8-M			4	1	3	
Communcation enhanced by NCEES	2-W			1	3	2	
Consistent expectations over time	1-W					1	
Totals		14-S	13-S	21-S	22-S	7-M	0-NR

Table 27 illustrates a strong frequency for the evaluation attributes, procedures and feedback, as well as teacher and evaluator attributes. All themes generated from comments were related to one or more of the attribute categories. No category produced a weak frequency, with context of the evaluation receiving a moderate strength coding.

Table 28

Themes and Attribute Categories that are Least Likely to Impact on Professional Growth

Themes – Elements of NCTEP least likely to impact growth	# of Times Mentioned	Connected to Attribute (Teacher, Evaluator, Procedures, Feedback, Context, Not Connected)					
		T	E	P	F	C	NC
Rubric/expectations (unclear; not aligned w/all aspects of instruction; lack of resources to meet requirements)	8-M			7		1	
Evaluator (bias;exaggerated ratings; little time to observe)	6-M			4	1	1	
Time-consuming	5-W			5	1	2	
Complications w/NCEES	3-W			2		1	
Lack of PD/practice/ collaboration	3-W		1	1		4	
I don't know	2-W						2
Computer-generated ratings/goals	1-W					2	
Totals		0-NR	1-W	19-S	2-W	11-S	2-W

Table 28 identifies the attribute categories of evaluation procedures and context of the evaluation with the highest strength codings for elements that are the least likely to contribute to professional growth. Attributes of evaluators and feedback produced weak frequency ratings. Two respondents commented that they did not know what components of NCTEP were least likely to impact growth, and the attributes of teachers

showed no relation as they were not connected to the themes.

Interviewees were asked if and how NCTEP had impacted their effectiveness as a teacher. Eight of the 14 (57.1%) teachers reported that NCTEP does impact their effectiveness. When asked to describe how their effectiveness was impacted, teachers in this reporting category most often stated that NCTEP encourages reflection (62.5%). Accountability (25%) and clarity of standards/expectations (12.5%) were also noted as contributing to the impact of NCTEP on effectiveness. Four of the 14 interviewees (28.6%) stated that there was somewhat of an impact of NCTEP on effectiveness, stating that accountability (50%); peer observations (25%); and rater bias (25%) were reasons for this impact. Two teachers (14.3%) stated that NCTEP had no impact on their teacher effectiveness, stating that they are more impacted by self-criticism (25%) and the use of student/parent and peer feedback (75%).

When asked if and how NCTEP impacted their attitudes toward teaching, seven of the 14 teachers (50%) shared that their attitudes were affected by the evaluation process. Approximately one third of these respondents (36.4%) explained that the process has challenged them to be better, more reflective teachers; however, an equal number of comments (36.4%) expressed concerns about the potential use of information from NCTEP and the pressure created by the process. The remaining teachers (27.2%) gave several ways that their attitudes were influenced, mentioning the amount of repetitiveness in the process, feeling frustration when professional development was not available for them to meet PDP goals, and concerns about observer bias. Fifty-percent of the 14 interviewees indicated that NCTEP had either no impact or a neutral impact on their attitudes toward teaching. Most (57.1%) stated that their love of teaching would not change with or without NCTEP. Slightly less than half of the comments (47.9%) gave

other reasons their attitudes were not affected, explaining that they viewed it as part of the job and use it as a feedback tool but that it does not change their attitude toward the teaching profession. One of the teachers added that her attitude might be different if the school culture or administrator was different.

Teachers in the follow-up interviews were asked to provide insight into several findings from the survey data. First, they commented on the lack of correlation between the attributes of teachers and the outcome ratings dealing with overall quality of the evaluation and the impact on attitudes, professional growth, and teacher effectiveness. Second, they considered the significance of the relationship between the context of the evaluation and the overall quality ratings and the impact on attitudes toward teaching, professional growth, and teacher effectiveness. Finally, teachers were requested to comment on survey results about the intended purpose of evaluation. Results had shown that slightly more than a third of respondents felt that there was a balance between accountability and growth. Less than a fourth (about 23%) of the respondens felt that the role was more toward accountability, while the majority of responses (about 41%) leaned more toward teacher growth. Six of the 14 teachers (42.8%) contributed additional comments to the interview protocol. Summaries of interviewee responses are following.

With regard to the lack of correlation of the attributes of teachers to outcome ratings, insights showed some commonalities. Three teachers (21.4%) stated that teacher attributes should play a part in the impact of NCTEP and were puzzled about this data. Two teachers (14.3%) speculated that teacher self-ratings were incorrect. One teacher (7.1%) wondered if the rating scale on the survey design could have contributed to this data. Attitudes about the evaluation instrument or evaluator were also discussed as reasons for this lack of correlation by four teachers (28.6%), and four other responses

considered whether teacher attributes were clearly addressed in NCTEP (28.6%).

Two primary themes emerged when interviewees were asked about the significance of the context of the evaluation to outcome ratings. Seven responses (50%) referred to professional development as key to the impact of NCTEP and reported this as a reason for the significance of this attribute. Five responses (35.8%) noted the connections of NCTEP to federal and state mandates and concerns about how evaluation data will be used in the future as an explanation for the correlation. Two responses (14.2%) were not connected to the context of the evaluation but commented on other attributes of NCTEP.

When discussing the survey results about the intended purpose of the evaluation, six of the 14 teachers (42.9%) commented that respondent answers may have been influenced by their administrator's approach to evaluation and school culture. Three teachers (21.4%) felt that the teacher's perspective on evaluation contributed to the perception of the intended purpose of evaluation; while one (7.1%) teacher stated that although the instrument was created for accountability, the way it is written promotes teacher growth. Four teachers (28.6%) answered in a manner that did not give a clear rationale for the disparity of survey responses.

Five interviewees chose to give additional comments for consideration in this study. One respondent explained that her husband's experience with teacher evaluation was very different from her own. Where her experience was very positive and enhanced her professional growth, his administrator used NCTEP in a punitive manner, leading him to compose multiple written responses to observation ratings. Another respondent discussed her participation in the Hope St. Project, which is an initiative developed by June Atkinson, North Carolina State Superintendent. She explained that the Hope St.

Project gathered feedback from teachers about education issues in North Carolina and how that data can be used to improve teaching conditions at the county and state levels. A third interviewee suggested that NCTEP should have more focus on being relational and give more opportunities for peer observations so that a team approach to education is fostered. Another teacher reported the need for more checks and balance in NCTEP, recommending that multiple evaluators conduct evaluations and contribute to the summative rating. The fifth teacher stated that he would like to see an open section on the evaluation rubric that allows principals to have the flexibility to address criteria that are not explicitly stated on the evaluation rubric. In his words, "it would be nice to inject a little bit more of humanity into the process."

Summary

The analysis of data presented in this chapter indicated that moderate, positive correlations exist between the overall quality of the evaluation and the attributes of evaluators ($r(187)=.204$, $p<.005$); evaluation procedures ($r(180)=.308$, $p<.001$); feedback ($r(180)=.308$, $p<.001$); and context of the evaluation ($r(180)=.349$, $p<.001$). Outcome ratings for measures related to teacher growth (changes to teacher behaviors and strategies) showed a moderate, positive correlation to attributes of evaluators ($r(184)=.161$, $p<.028$); evaluation procedures ($r(177)=.342$, $p<.001$); feedback ($r(177)=.342$, $p<.001$); and context of the evaluation ($r(177)=.468$, $p<.001$). Outcome measures related to teacher effectiveness (the impact of the teacher evaluation process on understanding of the teaching/learning process) were found to have a moderate, positive correlation to the attributes of evaluators ($r(186)=.200$, $p<.006$); evaluation procedures ($r(179)=.402$, $p<.001$); feedback ($r(179)=.402$, $p<.001$); and context of the evaluation ($r(179)=.471$, $p<.001$).

Additionally, a moderate, positive correlation was found between attitudes toward teaching and the attributes of evaluation procedures ($r(178)=.296$, $p<.001$); feedback ($r(178)=.296$, $p<.001$); and context of the evaluation ($r(178)=.371$, $p<.001$). A Pearson correlation was calculated that examined the relationship between the four outcome ratings and teacher attributes. A weak correlation that was not significant was found for teacher attributes and outcome ratings of overall ratings ($r(190)=.140$, $p>.05$); impact on teaching behaviors/strategies ($r(187)=.672$, $p>.05$); impact on understanding of teaching/learning process ($r(189)=.577$, $p>.05$); and impact on attitudes toward teaching ($r(188)=.315$, $p>.05$). Teacher attributes are not related to the outcome ratings as reported by the TEP (modified) survey.

For status ratings, teachers who received effective ratings showed positive, moderate correlations to the overall evaluation quality; impact on teacher behaviors/strategies; impact on understanding of the teaching/learning process; and impact on attitudes toward teaching in relationship to attributes of evaluation procedures, feedback, and context of the evaluation. Teachers who received highly effective ratings also showed positive, moderate relationships to attributes of evaluation procedures, feedback, and context of the evaluation as well as evaluator attributes.

Frequency data revealed that 78.7% of teachers chose a 4 or 5 on the Likert scale to describe themselves as teachers. These ratings were not normally distributed. Responses to the use of NCEES to enhance NCTEP were nearly equally divided between high (35%), mid (33%) and low (32%) rankings. When asked to identify the intended role of evaluation, results showed that slightly more than one-third (34.3%) reported a balance between accountability and growth. Less than one-fourth (23.7%) noted that the role leaned more toward accountability, while the majority of responses (42%) leaned

more toward teacher growth in their responses.

For teacher rankings of outcome ratings by experience levels, crosstabulations showed that approximately 39% of teachers with 20 years or more years of experience indicate a low ranking impact of the evaluation process on changes to their teaching strategies/behaviors, whereas 40% of teachers with 10 years or less experience give a high ranking for impact on teaching strategies/behaviors.

Analysis of crosstabs of the impact of teacher evaluation processes on the understanding of the teaching and learning process and respondent years of teaching experience show that teachers with 20 or more years of experience have the highest percentage of low rankings (51.6%) for this outcome rating, which is 18.8% higher than teachers in the 10 years or less group. Crosstabulations also showed that teachers with the most experience were the most likely (40.3%) to report a low impact on their attitudes toward teaching, whereas teachers with the least experience were mostly likely (45.9%) to report a high impact.

For teacher grade-level assignments, teachers in Grades 9 and above report the highest percentage (47.3%) of low rankings for the impact of teaching strategies/behaviors, demonstrating a 24.3% difference from Grades PreK-5 teachers; 23.8% difference from Grades 6-8 teachers; and a 27.3% difference to other teachers. Additional data show that 53% teachers in Grades 9 and above report a low impact ranking of the evaluation process on their understanding of the thinking/learning process, whereas 25% of this same group report a high ranking impact. Crosstabulations also indicate that a similar percentage of Grades PreK-5 teachers (35.9%) and Grades 6-8 teachers (35.3%) report a high ranking impact on the understanding of the teaching and learning process. A higher percentage of teachers (44.6%) in Grades 9 and above report

low rankings for the impact of the evaluation process on their attitudes toward teaching, which is approximately 15% to 25% higher than any other group.

Qualitative data collected through survey responses indicate a strong negative tone in comments about overall quality of NCTEP and the impact of NCTEP on teacher growth, teacher effectiveness, and attitudes toward teaching. A moderately negative tone was identified for the role of NCEES in enhancing NCTEP. Attributes of the evaluator and evaluation procedures had strong connections to the overall quality of NCTEP, whereas no attributes were strongly connected to the impact of NCTEP on growth. Evalaution procedure attributes were strongly connected to the impact of NCTEP on effectiveness as well as to the impact on attitudes toward teaching. Other dimensions to consider in further study of NCTEP that showed strong response frequencies were student factors and teacher factors, including relational qualities of the teacher.

Follow-up interviews showed strong response frequencies for the attributes of teacher, evaluator, evaluation procedures, and feedback as contributors to improved professional growth. The context of the evaluation had a strong response frequency in connection to elements that were least likely to improve professional growth. At least 50% of interviewees indicated that NCTEP does impact their effectiveness (57.1%) and their attitudes toward teaching (50%). They reported that NCTEP encourages reflection, provides accountability, and provides clarity of standards and expectations which improve effectiveness. Attitudes toward teaching were positively impacted by challenging teachers to become more reflective; however, the pressure created by the process and potential use of the information from NCTEP led to a negative impact. Slightly more than 57% stated that their attitude toward teaching would not change with or without NCTEP.

Chapter 5: Interpretations, Conclusions, and Recommendations

Introduction

Educators, the public, and politicians have a long-standing interest in improving the quality of education (Hechinger, 1978) and have viewed teacher evaluation as a means to that end (Darling-Hammond et al., 1983). This interest contributed to research advances in effective evaluation processes during the 1980s and 1990s at both the national and state levels (Bradshaw, 2002; Stacey et al., 1989). In North Carolina, TPAI was developed from this research base in conjunction with principal and teacher input and became one component of TPAS. Two primary purposes of TPAS existed: to enhance teacher growth through feedback and professional development and to provide summative ratings needed for personnel decisions (Stacey et al., 1989). TPAI is detailed in Appendix A.

TPAS was the prominent evaluation system used in North Carolina during the late 1980s and 1990s for beginning and nontenured teachers, although some North Carolina districts chose to develop local evaluation instruments for tenured teachers (Bradshaw, 2002). Mandates of the Excellent Schools Act of 1997, however, challenged locally developed evaluations and reinforced the validity of TPAI by requiring that districts using local evaluation alternatives demonstrate that their systems were at least as valid as TPAI (Bradshaw, 2002). A 1999 study utilized TEP (Stiggins & Duke, 1988),with modified questions applicable in North Carolina, to examine teacher and administrator perceptions of teacher evaluation in relationship to student learning, teacher growth, and school improvement (Bradshaw, 2002). Results of this study indicated that teacher satisfaction with TPAS was lower than desired; and while teacher ratings were generally high, teachers did not find the evaluation process useful (Bradshaw, 2002).

Revisions of the original TPAS occurred in TPAI-2000 (also, known as TPAI-revised). Changes to the system included the incorporation of performance standards related to student achievement and employee skills and knowledge (Flowers et al., 2000). Flowers et al. (2000) reported that this instrument was designed as a summative tool to defend against "recent attacks from legislators and the general public regarding the poor performance of some teachers" (p. 2). Districts were encouraged to develop and utilize formative evaluation systems in addition to TPAI-2000, stating that the ability for one system to meet both professional growth and accountability purposes was controversial (Flowers et al., 2000). TPAI-2000 served as the vehicle for evaluation in North Carolina until the current NCTEP was fully implemented in 2010-2011 (Public Schools of North Carolina, 2010).

The purpose of this study was to examine the impact of NCTEP on teacher perceptions of professional growth, teacher effectiveness, and attitudes toward teaching. The NCTEP manual states that its intended purpose is "to assess the teacher's performance in relation to the North Carolina Professional Teaching Standards and to design a plan for professional growth" (Public Schools of North Carolina, 2015, p. 4). Implications of findings are first considered in light of study samples from the two phases of data collection and then are organized by the research questions. Findings from the study may assist in determining if and how NCTEP is accomplishing its intended objective. As teacher evaluation is a highly political issue, the presentation of study findings, limitations of this study, and recommendations for further research may reflect implications for state and/or local evaluation policy.

Findings and Interpretations

The following section addresses findings related to the study samples in the two

phases of research.

Phase one: Online survey, TEP (modified)–study sample. Phase one of this study was conducted during a 2-week period from mid to late May 2016. This time frame was purposefully selected as end-of-year evaluations are often completed in this month, but end-of-year testing has not yet begun. Administrators were requested to introduce the opportunity to participate in this study to their teachers. Teachers were given the assurance that the survey results would not be connected to their school or administrators. All teachers in the district were invited to participate in the survey with a response rate averaging 22.5%.

A modified version of TEP, based on the TEP Questionnaire developed by Stiggins and Duke (1988), was used to gather data. TEP (modified) reflected the same key elements as the original TEP instrument and was designed with a similar purpose: to identify teacher perceptions of the overall evaluation process as it relates to teacher attitudes toward teaching, professional growth (behaviors), and teacher effectiveness (understandings). Revisions to the original TEP's directions and questionnaire items reflect NCTEP. When asking about the accuracy of teacher effectiveness ratings based on the summative teacher ratings (Standards 1-5) and on student growth measures (Standard 6), specific questions in the category of "Attributes of Evaluation Procedures" considered the six North Carolina Professional Teaching Standards (Public Schools of North Carolina, 2013b). Other changes to survey items asked for teacher experiences with content related to curriculum reforms and the Common Core State Standards and for teacher perceptions of the impact of political initiatives such as RttT on the evaluation process. Most survey items were Likert scale-type items, although some questions asked for nominal data. Opportunities were included for participants to add comments to

selected questions on TEP (modified). Providing commentary was optional. TEP (modified) is shown in full in Appendix F.

Through the original TEP, Stiggins and Duke (1988) validated the five key attributes of effective evaluations and utilized TEP to determine if these attributes were related to perceived teacher growth and improved teacher effectiveness. In this researcher's study, the five key attributes of effective evaluation also were examined in relation to four outcome rating questions. Attribute categories were (1) attributes of the teacher, (2) attributes of the person who observes and evaluates, (3) attributes of the procedures used to gather data on teacher performance, (4) attributes of the feedback, and (5) attributes of the evaluation context (Stiggins & Duke, 1988). Outcome rating questions asked for teacher perceptions of the impact of the evaluation process on the overall quality of the evaluation experience and the impact of the evaluation process on teacher understandings of the learning process (teacher effectiveness), teaching strategies and behaviors (professional growth), and attitudes toward teaching. According to Stiggins and Duke, the goal of their teacher evaluation research was to determine if the process of evaluation was meaningful to participants by enhancing their personal and professional growth. Similarly, this study intended to determine if teacher perceptions of NCTEP accomplished these outcomes.

At the time of this study, some teachers had received a status rating of in need of improvement, effective, or highly effective. This rating was based on EVAAS VAMs that were calculated after 3 years of data collection and evaluator summative ratings on Standards 1-5 (Public Schools of North Carolina, 2015). Teachers were asked to report their status ratings in the online survey. More than 92% of respondents reported either an effective (33.3%) or highly effective (58.9%) status rating. Only two respondents

reported an in need of improvement status; 10 reported that no status rating was given;

and three chose not to disclose their status ratings. Several factors may contribute to

these results. Since most of the respondents received satisfactory status ratings, it is

possible that these teachers felt more comfortable expressing their opinions about their

experiences with NCTEP and chose to participate in the online survey. A second

consideration for the high percentage of responses by effective and highly effective

teachers could be characteristic of the Widget Effect; which Weisberg et al. (2009)

explained is a deeply embedded, cultural phenomenon based on the assumption that

teachers are much like interchangeable parts. Weisberg et al. contended that one result of

the Widget Effect is an overstatement of high-performing teachers by evaluators. Studies

of the addition of growth measures into NCTEP report evidence of the Widget Effect,

noting that evaluator ratings of Standards 1-5 showed upward bias and "may limit the

potential of the evaluation system to provide a full range ofmeasurement and subsequent

formative assessment and feedback" (Lynn et al., 2013, p. 6).

Correlations between status ratings and outcome ratings demonstrated differences

in which key attributes of evaluation impacted effective teachers as compared to those

impacting highly effective teachers. Correlations for effective teachers showed stronger

relationships for evaluation procedure, feedback, and context of the evaluation attributes

than correlations for the highly effective teachers. Outcome ratings for highly effective

teachers demonstrated weak correlations to evaluator attributes, whereas no relationship

was found for effective teachers. Neither group of respondents indicated that teacher

attributes were related to outcome ratings for overall quality of NCTEP or the impact of

NCTEP on teacher growth, effectiveness, or attitudes toward teaching as was the case for

overall survey responses.

Participants in the follow-up interview phase of this study were asked to explain why a correlation may not have existed between teacher attributes and the outcome ratings. Three teachers expressed bewilderment at this lack of correlation. Teacher 3 commented, "I would say that teacher attributes would play a huge part in how they view the evaluation process, in my opinion, because if you're flexible as a teacher, then you can kind of receive feedback a little bit more." Teacher 66 remarked,

I don't think the evaluation checks for things like, teacher attributes–flexibility in the classroom, are you organized, are you efficient–I don't think the evaluation–it looks more at content and classroom management and teacher leadership–those kinds of things. So, how you see yourself as a person, which would be your teacher attributes don't necessarily show up on that evaluation tool.

Teacher 12 stated, "That didn't show? Hm. Well. It could mean that, maybe people didn't rank them like–like they really felt . . . I would think that would have to be it. Something wasn't rated right because it ought to show." The validity of this opinion may be supported through the analysis of frequency data which revealed that 78.7% of teachers chose a 4 or 5 on the Likert scale as their response to teacher attribute questions. While this may not necessarily mean that "something wasn't rated right," it does demonstrate an uneven distribution of responses to this category of attributes.

Information about teachers' years of teaching experience and the grade level taught at the time of their last summative evaluation was collected through the online survey. Teaching experience data showed that approximately 32% of teachers had 10 or fewer years of experience; 35% had 11-19 years; and 33% had 20 or more years in the profession, reflecting evenly distributed samples. Respondents' grade levels taught indicated that approximately 28% were Grades PreK-5 teachers; 27% were Grades 6-8

teachers; 29% taught Grades 9 and above; and 16% fell into the other category of certified teachers. These percentages are similar to the grade-level distributions of certified teachers in this district. District reports categorize approximately 48% of teachers are assigned to Grades PreK-8; 24% teach in Grades 9 and above; and 27% are designated as other. A lower response rate from teachers in the other category may have been received as this group includes school counselors, psychologists, and media coordinators among other certified teachers. Some of these individuals may have been eliminated in the survey's qualifying question as their evaluations have only recently been added to NCEES, and they may not have considered this study applicable to their positions. Data indicate that teaching experience and grade-level assignments may influence teacher perceptions of NCTEP and will be described in further detail in relation to the connected research question.

Narrative responses were also returned for six questions in the online survey. The percentage of participants who provided additional narratives to survey questions varied between 8.2% and 21.2% of the total survey completions. Overwhelmingly, these responses projected a negative tone. This high percentage of negative comments seemed inconsistent with overall quality outcome ratings from the survey, in which approximately half of the respondents rated the overall quality of the evaluation positively. The assurance of anonymity may have contributed to the tone of comments in the online survey as teachers may have felt comfortable being frank in discussing their experiences with NCTEP. The timing of the survey also may have contributed to this negativity. In one follow-up interview, for instance, Teacher 10 noted,

Well, I will tell you, I'm going to be honest with you about the survey. You sent the survey out after I had been observed and, uh, it was before my summative, but

> it was after the last time I was evaluated and, I mean, I didn't receive a bad
>
> evaluation–that was not the issue–I have never received a bad evaluation–but,
>
> every time I'm evaluated with that rubric, I just get so hostile. . . . I felt really bad
>
> afterwards, I thought, "Well, I've already sent it."

This teacher explained a primary reason for frustration was that elements of the rubric

were not relevant to the characteristics of all students.

Thirty-eight comments were submitted to the online survey question that asked

for additional topics that should be added to the study, accounting for 21.2% of the total

responses. Themes that emerged from these comments showed that teachers favored

consideration of student factors such as motivation, attendance, behavior, and intellectual

qualities in the evaluation process. One teacher stated, "Motivation of children whether

from self, parents, or teacher is not included. Every bit whether negative or positive falls

back on the teacher which does not seem realizable (sic) or fair." Teacher factors,

including relational qualities and personal characteristics, were also noted multiple times.

These characteristics were described in one comment as the "ability to work with others

as a team, [as well as] how cooperative and flexible we are working toward our vision as

a school." Another comment noted that "consistency and reliability as an employee"

were dimensions that should be considered in this study. One comment summized

multiple themes, stating,

> The teacher evaluation does not take into account the inconsistent availability of
>
> resources, various degrees of academic achievement below grade level in each
>
> classroom, parental involvement, parental control of the school environment,
>
> student motivation, behavior climate, [and] limits on actions taken by teachers
>
> regarding rewards and consequences for behaviors related to academics or school

environment.

Additional comments provided by a teacher in the follow-up interview phase of this study echoed these sentiments and offered a suggestion for addressing these factors on the evaluation instrument, stating that "an open section for principals to address criteria that are not explicitly on the evaluation" would be a way to "inject a little bit more of humanity into the process. Not just paper–pencil–click, click, click." Comments such as these imply that teachers perceive NCTEP does not fully consider all factors impacting their performance.

Phase two: Follow-up interviews–survey samples. Phase two of this study involved 14 teachers who volunteered to participate in face-to-face or phone interviews during July 2016. Teaching experience of interviewees ranged from 6 to 27 years. Five teachers reported 0-10 years of teaching experience; four had 11-19 years of experience; and five teachers had 20+ years of experience. Three of the teachers were assigned to Grades PreK-5; three worked in Grades 6-8; seven worked in Grades 9-12; and one was considered in the other category, which has been previously described. This sampling represented a smaller but similar representation of these groups in comparison to the online survey sample with the participation rate of teachers in Grades 9 and above being slighty above those in Grades PreK-8.

Teacher experiences with NCTEP were diverse. While all had been observed, five had served as peer observers. Teachers who had been peer observers discussed the value of opportunities NCTEP created for peer conversations. One peer observer commented that he thought "the best thing about [the NCTEP] is the peer observer component," reasoning that

If my peer comes in and says . . . "did you think about this," I'm probably going

to think about that a little bit more than if an administrator came in. That may be the simple difference just between, you know, the connection that I have with a peer versus an administrator or evaluator.

Three participants discussed their experiences with TPAS that was used prior to NCTEP and expressed varying perceptions of the two systems. One interviewee stated that "when we had the previous teaching evaluation instrument–the one before the– whatever it is that we're using now–I feel like I got better feedback from evaluators, using the old system." On the other hand, a different teacher described NCTEP as "a little bit more clear of what they're looking for" and discussed

> a time where, I guess when I first started teaching–that would have been mid-
> 90s–the evaluation was basically, "Here's what I saw, good job," there wasn't a
> whole lot of category to it, there wasn't a whole lot of knowledge, on my end, of
> what they were even looking for . . . I think there's more direction with the
> current system.

One teacher brought a distinct prespective to the interviews as she reported,

> I was part of the original team that helped design the evaluation process for the
> state. So, I was part of breakout groups, regional here, in the Northwest region,
> where we talked aobut what we were looking for–it was, uh, basically a question
> and answer session that we did as a region. And then, I was pulled in to, by the
> NCAE to be part of a focus group of approximately 30 teachers to help work with
> the group out of Colorado to design the original document.

She continued by explaining that

> we were trying to get away from the old evaluation, which was a checklist based
> on the seven point lesson plan . . . where you could talk about work that you, uh,

have done on your own, or things you have done on the state or national level . . .

so that was the intent of the document.

She stated that the current NCTEP uses components of the document she worked on but that the format of the rubric shows the exemplars that the breakout groups developed "super-imposed onto a checklist." From her perspective, "its kind of two steps forward, three steps back–is what's happened."

Another unique experience reported by a teacher during the follow-up interviews was his work with ASW, which is the growth measure for educators in content areas where student artifacts are used to determine growth. He expressed strong concerns about the integrity of this system stating that

there are so many ways on that particular system to be dishonest that I think it's a

royal waste of time and the state's money . . . it is not just a legitimate system

from top to bottom. There are too many ways to wiggle out from under the rules.

As for ASW's impact on teacher growth, this interviewee commented, "It's not really a system where you can get feedback to make you a better teacher. It, uh, it's really documentation on top of documentation."

The vast range of experiences described by teachers who volunteered to participate in follow-up interviews demonstrated a high level of understanding of and reflection on teacher evaluation processes in North Carolina. Many of them had positive experiences to report, but some described negative experiences typically involving relationships with evaluators. When thanked for participating in the survey, one teacher said, "It's important work." To these participants, NCTEP is a significant topic for discussion.

Research Question 1. What attributes of NCTEP impact teacher perceptions

of professional growth? Correlation data, crosstabulations, survey comments, and follow-up interviews provided insight into what evaluation attributes teachers perceived as impacting professional growth. Correlation data indicated that the attributes of the evaluator, evaluation procedures, feedback, and context of the evaluation had weak to moderate relationships with teacher perspectives of changes in their teaching behaviors and strategies, which are indicators of teacher growth. When disaggregated by status ratings, relationships were identified between evaluation procedures, feedback, and context of the evaluation for both effective and highly effective teachers. Teacher and evaluator attributes did not reflect significant relationships with growth outcome ratings for either of these groups.

Crosstabulations showed that the impact of NCTEP on changes in teaching behaviors/strategies differed based on amount of respondent teaching experience. Teachers with 20 or more years of experience reported lower ranking impact outcome ratings than teachers with less experience. Teachers with 10 years or fewer of experience reported the highest impact of the three groups, and teachers with 11-19 years of teaching demonstrated the highest impact rankings in the mid ranking range. Since teacher evaluation in North Carolina has undergone multiple changes since the late 1990s, it is likely that teachers with higher experience levels are less likely to view the process as having a consistent impact on their professional growth.

Crosstabulations based on grade levels taught and professional growth also indicated that teachers perceived the impact of NCTEP on growth differently based on their teaching assignments. Most notably, almost half of the teachers in Grades 9 and above reported low rankings for the impact of NCTEP on growth. On the other hand, 40% of teachers in the other category reported high rankings of NCTEP's impact.

Teachers in Grades PreK-5 and Grades 6-8 had similar impact rankings with a higher percentage of mid and high rankings reported than for low rankings. Comments collected through the online survey and in follow-up interviews gave little explanation for this occurrence; however, there is some evidence that perceptions of teachers in the other category may have been influenced by their experiences with the ASW growth measure.

Comments from the online survey indicated that overall teachers perceived much of the impact of NCTEP on changes to teaching behaviors and strategies to be negative. They frequently noted that their professional growth was influenced by factors beyond the evaluation process, referring to the role they play in their own professional growth. For instance, one teacher commented,

> I am a motivated teacher who wants to improve myself so I have studied the evaluation standards and used them as a way to improve myself in areas that I would not have thought of. However, the PROCESS itself is frustrating because evaluators are not using the process as a way to improve teaching at my school– it's just a matter of "checking the right boxes."

Although the attributes of evaluators did not show significant relationship in the data analysis, comments such as these indicated otherwise. Teachers mentioned their interactions with evaluators and noted that feedback they received contributed to professional growth but that other elements of NCTEP were not useful. The following statement illustrates this clearly.

> My evaluator's comments help me see my strengths and weaknesses which I appreciate. However, the EVAAS value-added score is not helpful at all. I exceeded growth but got no feedback as to why. I just get the score. I get no info about why or how to use the info to inform my decisions about the next year.

>   And besides, by the time I get the EVAAS score, I am already–almost ready–to
>
>   give the exam again. Useless!!

ASW, the growth measure used to calculate EVAAS for selected teachers, received
commentary in its connection to teacher growth.

>   ASW–Negative impact. I teach Spanish and it is hard to prove on paper that
>
>   students don't know the content before starting a unit of study. It is ALL new to
>
>   them. I have to take valuable class time to give them assessments on material
>
>   they have never seen.

At least one other teacher also felt that the time required for NCTEP was "a waste."

Teacher attributes generated only a weak correlation with the outcome rating for
changes in teaching behaviors. Comments revealed, however, that teachers do feel that
personal characteristics such as motivation and work ethic play a key role in their growth
as an educator. They described themselves as "a motivated teacher," "a reflective
practitioner," "continually changing," and "a good teacher"; and they stated, " I do good
work," "teach what I am supposed to teach," and "wake up everyday with intentions of
doing my best!" Given these statements, the lack of correlation was puzzling and was
explored further in follow-up interviews. Interviews revealed that teachers were
perplexed about this correlation.

Follow-up interviews provided definitive responses to the research question,
"What attributes of NCTEP impact teacher perceptions of professional growth?"
Participant remarks identified the quality of feedback/productive conversations as the
strongest element to impact professional growth. One teacher stated that NCTEP gives
"kind of a jumping off point for some good, rich conversation." The ability of NCTEP to
promote self-reflection and goal setting was also mentioned frequently by respondents

and was synthesized in statements about the impact of other attributes as the following

comment illustrates. In response to a question that asked teachers to identify which of the

five key attributes was most significant to professional growth, one teacher replied,

> I would say feedback. Because if you don't know what they observed–if you
>
> can't reflect on that and look at what it looked like through others' eyes, then that
>
> makes it more difficult for you to grow. So, being able to look at what others are
>
> seeing in your classroom, getting feedback from them, and being able to reflect on
>
> that and–and taking that to see how you can adapt your own teaching and skills in
>
> the classroom.

Other elements that demonstrated moderate frequencies to positive impacts on

professional growth were also identified. Interviewees noted that the standards and

expectations identified through NCTEP provided a focus for professional growth, stating

that the most helpful aspect of NCTEP in improving professional practice was "being

able to identify what standards, what specifically is being looked for and valued, and also

being able to have a clear definition of, in each category, what it takes to be accomplished

or proficient." PDP and related professional development were mentioned as another

element that provided focus for professional growth. Discussion of PDP revealed that

this element had positive impacts on interviewee teaching practices. Confidence in the

evaluator also influenced how NCTEP was seen as contributing to professional growth.

One teacher's response exemplified both of these components as well as alluding to the

importance of feedback. She stated that the most helpful aspect of NCTEP for improving

professional practice was PDP, explaining,

> Just making sure that you follow through with your PDP and, if there's something
>
> that you need to work on, you know, that's a good time for them [the observer] to

let you know–or if you're doing something good that you need to continue doing, they'll let you know that during the interview process.

Due to the intricacies of responses such as the one above, when data were analyzed to make connections between themes and their correspondence to key attributes, strong, positive connections were shown between attributes of teachers, evaluator, evaluation procedures, and feedback. Context of the evaluation was moderately connected to teacher perceptions of the positive impact of NCTEP on their professional growth but received a stronger frequency rating, an attribute that was perceived as having an ineffective impact on professional growth. Attributes of the evaluation procedures also demonstrated a strong frequency rating as an attribute that was perceived as having an ineffective impact on professional growth. These ratings imply that while it is possible to identify what attributes of evaluation impact teacher perceptions of professional growth, discerning why these attributes have both positive and negative impacts needs to be examined in more detail.

Research Question 2. How do the attributes of NCTEP impact teacher perceptions of teacher effectiveness? To determine how the attributes of NCTEP impact teacher perceptions of teacher effectiveness, crosstabulations, survey comments, and follow-up interviews were analyzed and synthesized. Crosstabulations of years of teaching experience by the impact on teacher effectiveness rankings indicated similar results as those demonstrated for the impact on teacher growth. Teachers with less experience, 10 or fewer years, reported higher ranking impact (37.7%) than any other group; and teachers with 20 or more years of experience reported the lowest ranking impact (51.6%) than any other group. Teachers with 11-19 years of experience had the highest percentage of mid-ranking impact (44.1%) ratings. Grade-level assignment data

are also similar to those reported for teacher growth. Teachers in Grades 9 and above report lower impact rankings (53.6%) for teacher effectiveness, whereas 40% of teachers in the other category give a mid ranking and 33.3% have a high ranking for NCTEP's impact on effectiveness. Grades PreK-5 and Grades 6-8 teachers vary less than 15 percentage points, between 27-40% across all ranking categories. These findings imply that teacher perceptions of the impact of NCTEP on teacher effectiveness (their understanding of the teaching and learning process) and teacher growth (changes in teaching behaviors/strategies) have an association with their years of teaching experience and the grade levels taught during their last completed evaluation cycle. Additional investigations into these demographics, which are not included in this study, may identify reasons for these commonalities.

Two themes were generated with strong frequency rates from additional comments concerning teacher effectiveness in the TEP (modified) survey. These were (1) that NCTEP had little or no impact on their teacher effectiveness and (2) evaluation procedures impacted their effectiveness. With the exception of evaluation procedures, all other attributes received weak to moderate frequency rates for this topic. While teachers were not greatly divided on the point that evaluation procedures was the attribute that was most connected to teacher effectiveness, they were less in agreement about whether the impact was positive or negative. Overall comments were negative, but five of the 14 remarks were interpreted as having both a positive and negative tone or as being neutral. Two comments were entirely positive. This may imply teachers view certain components of NCTEP as helpful but feel that there are some procedural changes that would enhance the process. Comments from the survey and follow-up interviews support this possibility.

Survey comments related to teacher effectiveness exemplified the diversity of perspectives about evaluation procedures and illustrate how the components of NCTEP were viewed. For instance, one teacher remarked,

> The evaluation process itself was completely useless in terms of helping me become a better teacher. I improved, with time, by working with brilliant professionals in my department who actually cared about my well-being and professional career. The sole exception to this was this year's peer evaluation post-observation conference.

In contrast, another teacher reported that due to NCTEP, she

> began to see a correlation in the types of lesson methodologies and how they impacted student growth and academic success. I also began to cull out lessons that were not beneficial, adapt lessons in order for students to better understand the concept, and create new lessons that were more rigorous.

Some comments were blunt, declaring with no explanation, "The evaluation process has nothing to do with my understanding of the teaching process." While others saw the potential in the process, explaining, "If used correctly, perhaps it would." One teacher composed a lengthy but insightful comment about the positive and negative aspects of NCTEP, noting,

> There is some impact with the evaluation process. I like challenging myself to grow. I like setting goals that help me focus on one or two areas. I do not like having the same goals year after year. I like reflecting on my practice, which the end of the year evaluation accomplishes. Overall, the length of the process takes away from its effectiveness. With so much else to do, especially at the end of the year, I know teachers who just check off boxes because they know ultimately it's

the principal's decision. I do this for myself because I want to know if I fell short in any area so that I can improve the following year. Last year, I honestly did not spend as many hours on my final evaluation. It's too long, and we have too much other paperwork and duties required.

Some teachers described factors they felt were more important to teacher effectiveness than NCTEP, stating, "My understanding is influenced much more substantially by research and current documentation about teaching and learning that is available online and from outside sources"; and "It is my constant reflection based on continuing education–reading, contact with other teachers, research–that have a strong, lasting impact."

Follow-up interviews were as divided in their perceptions as the survey comments. A majority of the responses (57.1%) affirmed that NCTEP did impact their effectiveness. These teachers identified the reflection and accountability aspects of the process, in addition to clarity of the standards and expectations, as contributing to the impact. One teacher commented that he appreciates "the consistency of the questions from year to year and then being able to see how I grade myself versus how someone else would evaluate me." Four of the 14 interviewees (28.6%) stated that there was somewhat of an impact of NCTEP on effectiveness. Accountability, peer observations, and rater bias were indicated as reasons for positive and negative impacts on teacher effectiveness. Speaking to the accountability aspect, a teacher commented, "It keeps us in check . . . if you want to excel . . . it gives you something to understand more about yourself and figure out why and, 'is this something beyond my control or something I need to fix.'" Two teachers (14.3%) stated that NCTEP had no impact on their teacher effectivenss, stating that they are more impacted by self-criticism (25%) and the use of student/parent

and peer feedback (75%).

Teacher effectiveness and its relationship to accountability were frequently mentioned throughout the follow-up interviews. Overall, this group of participants welcomed the opportunity to be evaluated as long as the process was relevant to improving the educational experiences of their students. As one interviewee stated, "it is important that we are evaluated, but it needs to be pertinent . . . I do want to have an effect on kids." Another commented,

I feel like I want to score high on the evaluation . . . I feel like I'm letting the school down if I don't . . . I didn't ever want to feel like an inadequate teacher . . . I'm there for the kids. I'm there to help them and so, if I'm not meeting that . . . then I'm failing them, so that's kind of how I view it.

These comments are indicative of the responsibility that teachers feel toward their students. Components of NCTEP, including evaluation procedures, that are not perceived as having a positive impact on students may be viewed negatively by teachers. As these follow-up interviewees demonstrated, teachers are concerned about how NCTEP is relevant in helping them become more effective teachers for their school and its students.

Research Question 3. How does NCTEP impact teacher attitudes toward teaching? Crosstabulations, survey comments, and follow-up interviews were instrumental in investigating how teacher attitudes toward teaching were impacted by NCTEP. Continuing the pattern established with teacher growth and effectiveness, crosstabulations between the impact of NCTEP and attitudes toward teaching documented that teachers with the most experience were the most likely (40.3%) to report that NCTEP had a low impact on their attitudes toward teaching. Teachers with the least

experience were most likely (45.9%) to report a high impact. For high ranking responses, there is an 8.8% difference between these two groups. Low ranking responses showed a greater difference (20.6%) between experience groups. Commentary from a teacher with 22 years of teaching experience provided insight into why this pattern may exist and gave some perspective into how less-experienced teachers may be impacted by NCTEP. He stated,

> In a general sense today, as a 22-year teacher, they're going to be very few things that an administrator can tell me that's going to be helpful from the classroom management situation or organizational standpoint. Partially because I am what I am at this point. My teaching's not going to change a lot in the next seven years. I'm pretty rooted. I know what works for me and I know what works for my students. I know where we can go–there are times when I think of some administrators who try to inject their personal wants and thoughts into, especially young teachers, to a level that's maybe too much. They will tell them that they have to do something a certain way. And, as a–as a teacher, I know that there's lots of ways to make the trip from here to there–I may not all be in a straight line– some of them might have, uh, times when you're in reverse because you've got to back up and punt, but, I–I think sometimes we–we drive, especially, traditionally core tested subjects to all teach the same way–not sure it all really pans out that well.

If this is a common attitude among more experienced teachers, this way of thinking can contribute to these teachers' perceptions of NCTEP's impact on their attitudes toward teaching. A closer look at the differences in the perceptions of teachers with varying experience levels is warranted but is not included in this study.

Crosstabulations for NCTEP's impact on attitudes toward teaching showed some variances from patterns for NCTEP's impact on teacher growth and effectiveness. A higher percentage (44.6%) of teachers in Grades 9 and above still report low rankings for the impact of the evaluation process on attitudes, which is approximately 19% to 25% higher than any other grade-level group; however, teachers of Grades 6-8 reported the highest impact ratings for attitude (49%) and are nine percentage points higher than teachers in the other grade level category (40%), 11.1% higher than Grades PreK-5 teachers (38.9%), and 18.6% higher than Grades 9 and above teachers (30.4%). Research in this study cannot account for the marked difference in NCTEP's impact on Grades 6-8 teacher attitudes toward teaching.

Survey comments related to attitudes toward teaching were strongly, but not solely, negative. All themes identified as impacting teacher attitudes had moderate to weak frequency ratings with the exception of the category, "has no impact/little connection to teaching." Moderate themes were categorized as increased stress, busy work/one more thing to do, and influenced by the administrator. When themes were classified by attributes of the evaluation, evaluator attributes received a strong frequency rating of 13. Context of the evaluation received a frequency rating of eight, which is moderate strength. Feedback (5) and teacher (2) attribute frequencies were weak, and feedback (0) was not related to NCTEP's impact on attitudes toward teaching.

Comments provided by survey respondents indicated that if NCTEP had an impact on attitudes toward teaching at all, it was perceived to be a negative impact. Only four comments referenced positive outcomes, noting that NCTEP encouraged self-reflection and that good evaluators contribute to attitudes toward teaching. Several comments expressed deeply negative sentiments about the impact NCTEP had on their

attitudes toward the profession. One teacher said, "made me want to quit teaching and left me with no confidence with myself even though the previous year's evaluation was great, this year's was a total bomb even by the same observer." Another teacher stated, "I am at the end of my career and if I were a new teacher I would seriously consider leaving the profession," while yet another described NCTEP as "the single most stressful part ofmy teaching career," calling it "demeaning, demoralizing, and stressful in the extreme." Additional comments were less emotional; but teachers still expressed frustration with elements of NCTEP, pointing to NCTEP's lack of "connection to teaching," describing the process as "busy work."

The evaluator was identified as a factor in NCTEP's impact on attitudes toward teaching in multiple comments. One teacher explained, "I do not believe the process in and of itself generates any impact. Again, it is the professionalism and effectiveness of the evaluator, not the tool, that determines its impact." Yet, another teacher stated, "It frustrates me because I don't feel like the people doing the evalautions take it as seriously as they should." Others expressed a lack of confidence in the process, stating NCTEP "can never actually see how well a teacher teaches," and explaining, "I see it as just another bureaucratic hurdle to cross." One comment clearly reflected the teacher's focus on the students rather than evaluation. He stated, "I hope the evaluations would not be a change agent in how I feel about my attitudes towards (sic) teaching. Hopefully, my students will continually be what motivates my attitudes about teaching."

Although the survey comments concerning the impact on teacher attitudes toward teaching were primarily negative, survey responses on the Likert scale items were more balanced with 32% indicating low rankings, 30.4% at mid ranking, and 37.6% giving a strong impact ranking. In follow-up interviews, seven of the 14 teachers (50%) shared

that their attitudes were affected by NCTEP. Equal percentages of these teachers

reported that the process challenged them to be better, more reflective teachers (36.4%)

and that they were concerned about the potential use of information from NCTEP and the

pressure created by the process. As one teacher noted, "it's made me a much, much

stronger teacher as far as reflecting on–the reflection piece–you know, reflecting on what

worked, what didn't work, how can I change to make it better, and using all of that

feedback." Concern about how the information would be used and the pressure created

by the process was summed up in these statements from one teacher:

> One big negative that comes to mind . . . is some of the ways in which data is used
>
> against teachers, does not take in poverty and does not take in growth . . . the
>
> standard does reflect what I teach–but, the end of grade scores do not necessarily
>
> . . . there's what I'm saying about–especially schools that are low performing–if
>
> they're not evaluating how much growth is–is made, then that's not a fair
>
> comparison.

According to Stiggins and Duke (1988), the context of the evaluation is indicative of the

environment in which the evaluation occurs. The manner in which information from

NCTEP is used is contingent upon the attribute of context of the evaluation. The

remaining teachers (27.2%) who remarked that NCTEP impacted their attitudes toward

teaching mentioned the amount of repetitiveness in the process and feeling frustration

when professional development was not available for them to meet PDP goals and

concerns about observer bias. These themes are closely connected to the attribute of

context of the evaluation.

In correlation data, the context of the evaluation and attitudes toward teaching

demonstrated a moderate, positive relationship; therefore, additional data were analyzed

to investigate reasons for this connection. Likert scale items from TEP (modified) that related to the evaluation's context considered (1) whether NCEES enhanced NCTEP, (2) if professional development was provided to support PDP goals, (3) the intended role of evaluation, and (4) the impact of state and federal mandates on NCTEP. The online survey generated data about the use of the NCEES online platform to enhance NCTEP. Overall, responses indicated that 31.9% of respondents felt that NCEES was not an enhancement to the evaluation process; 33% were neutral; and 35.1% stated that it did enhance NCTEP. Survey comments were less balanced with nine of the 15 comments being negative and 12 of these stating that NCEES did not enhance the process, was a waste of time, or was confusing. Since this question was an element of the context of the evaluation attribute, these perspectives may be related to the impact of NCTEP on attitudes toward teaching.

A question was included in the follow-up interviews to further explore the relationship between context of the evaluation and teacher perceptions of the impact of NCTEP. When asked to explain why this relationship may exist, interviewees discussed the need for professional development in connection with their PDPs and reasons why it does not exist. They also referenced the influence of school and county goals on PDP and remarked that NCTEP was "tied to Race to the Top and No Child Left Behind–or whatever its current incarnation is." However, one teacher gave a unique perspective on the significance of the context of the evaluation, stating,

> What my gut reaction to that is that, I think, people see the potential in it, you know, and, I think, that people, that there is the possibility of it truly driving practice and their effectiveness. I'm guessing people just aren't seeing that, in reality.

The impact of NCTEP on attitudes toward teaching should be an area of concern due to the highly negative comments. This teacher's statement identifies a key issue. Although NCTEP has the potential of improving teacher growth and effectiveness, it is not reaching this potential when the process is put into practice. Further exploration of NCTEP in practice should be considered to assess this teacher's theory.

      Intended purpose of the evaluation. Assessment of a teacher's performance based on the expectations of the North Carolina Professional Teaching Standards and designing a plan for professional growth is the intended purpose of NCTEP (Public Schools of North Carolina, 2015). Eight goals of the evaluation process support this purpose. The goals are to

> 1) serve as a measurement of performance for individual teachers; 2) serve as a guide for teachers as they reflect upon and improve their effectiveness; 3) serve as the basis for instructional improvement; 4) focus on the goals and objectives of schools and districts as they support, monitor, and evaluate their teachers; 5) guide professional development programs for teachers; 6) serve as a tool in developing coaching and mentoring programs for teachers; 7) enhance the implementation of the approved curriculum; and 8) inform higher education institutions as they develop the content and requirements for teacher training programs. (Public Schools of North Carolina, 2015, p. 5)

While not explicitly researched in this study, data analysis revealed teacher perspectives as to if, and how well, NCTEP is accomplishing its intended purpose and goals. For instance, the intended purpose of NCTEP was viewed as balanced between accountability and growth by slightly more than one third (34.3%) of the responsdents to TEP (modified). Less than a fourth (about 23%) of the respondents felt that the role was more

toward accountability, while a majority of responses (about 41%) leaned more toward teacher growth. This finding demonstrates that the instrument is perceived by some teachers as meeting the dual purposes of accountability and teacher growth as it is designed to do, but supports concerns by Marzano (2012) and Flowers et al. (2000) as to whether a single instrument can meet both purposes.

To further explore this topic and examine the significance of the context of the evaluation attribute on teacher perceptions of NCTEP, interviewees were asked to comment on the distribution of responses from the online survey. One teacher stated, "I'm glad it was kind of balanced there"; and another said, "the majority of us feel like the intent of evaluation is to build teachers to be better educators, but, you know, you've got the angry ones out there." One rationale for the distribution was attributed to the evaluator by one teacher who commented, "I think that goes back to the use of the tool, and how well that administrator uses it . . . so, I can understand why it runs the gamut."

Contrasting opinions were illustrated in comments from two teachers and give rationales for why the purpose should be more accountability or more growth-focused. Speaking to accountability, one interviewee stated,

I think the instrument was created so that we'd have some kind of accountability, but because of the way it's written, it's to help us grow. So, I would see–I would think that it would be a mixture of both. So, you know, we have to, you know, in everything that we do–in every job out there, there has to be some sort of accountability and, so this is our accountability. And then, when we have accountability, we have to make sure that we're rising to that–that standard. And if we don't possess that, then we need to go out and figure out how to obtain that accountability.

Speaking from the growth perspective, this teacher commented,

> I would personally like to see it be more about teacher growth than about teacher
> accountability, because if we're doing what we're supposed to do, we're only
> going to make ourselves and the kids that we work with better. If it has to come
> down to teacher accountability, it would seem like, you know, somebody's either
> not doing what they need to do or they're waiting on someone to tell them how
> they need to do things better. But I would like to see it more about the growth
> side and making sure that we're already doing good things and what can we do to
> make it better.

Interviewee comments supported the distribution of responses found in data from the
online survey. Teachers in this study perceive that NCTEP accomplishes dual functions.
As follow-up interview comments indicate, teachers differ, however, in what the primary
purpose of NCTEP should be and are concerned about the evaluator's role in the process.

Limitations

The limitations of this study include those primarily associated with quantitative
studies, survey samplings, and self-reported measures.

Quantitative design. In this study, the quantitative design utilized regression
analysis of the four outcome rating questions in relation to the five key attributes of
effective evaluation as determined by TEP (modified), which was based on Stiggins and
Duke's (1988) earlier TEP questionnaire. While this quantitative methodology identified
relationships between the outcome ratings and the attributes of the evaluation, the
quantitative design allowed only for predictive relationships and did not identify casual
relationships. The collection of qualitative data was necessary to support the findings
generated from quantitative analysis and to answer the research questions.

Survey samplings. This study was limited by the survey samplings. Only teachers who had completed NCTEP, including a summative evaluation, were considered in this research. Teachers who had not completed the full NCTEP were disqualified from participating in the online survey. Since participants for the follow-up interviews were identified from the survey, teachers who had not completed the full NCTEP were also excluded from this phase of the study. Survey results indicated that the majority of respondents to the survey received effective or highly effective status ratings. The perspectives of teachers in these categories may differ from those receiving in need of improvement ratings. This factor may limit the study. Finally, this study was limited to one midsized school district in Western North Carolina, and this may limit the generalizability of the results.

Self-reported measures. As this study used a survey, validity of the results depended on the accuracy of teacher self-reporting. Teachers' potentially limited or biased remembrances of their evaluation experiences may have impacted the quality of the data collected. The time of year when the survey occurred may also have limited this study as teachers were under higher stress near the end of the school year. The lack of inclusion of evaluator perspectives in relation to teacher self-reports limited the analysis of the accuracy of teacher self-reporting.

Conclusions

Most respondents to this survey self-reported that they received status ratings of effective or highly effective. Overall status ratings are based on EVAAS student growth data and evaluator rankings of Standards 1-5 (Public Schools of North Carolina, 2015). Two conclusions are based on these results. One, it seems likely that teachers who receive satisfactory ratings in NCTEP are willing to voluntarily participate in research

about the evaluation process. Responses from interviewees and from survey comments support this conclusion as participants described themselves as "a motivated teacher" and "a reflective practitioner," stating that "I don't want to be just standard." A second conclusion drawn from these results is that the high number of effective or highly effective teachers (92%) responding to this study demonstrates that the Widget Effect (Weisberg et al., 2009) is impacting teacher ratings. Weisberg et al. (2009) contended that one result of the Widget Effect is an overstatement of high-performing teachers by evaluators. The Widget Effect has been previously noted in studies of NCTEP (Lynn et al., 2013) and may be evident in this study as well.

In this study, survey data showed that there was not a significant relationship between teacher attributes and teacher ratings of the quality of the evaluation process or its impact on attitudes toward teaching, professional growth, or teacher effectiveness. These results are similar to those obtained by Stiggins and Duke (1988) when they initially administered the TEP Questionnaire, which linked only one question within the teacher attribute category to outcome ratings. When interviewees in this study were asked to comment about the lack of relationship between teacher attributes and outcome ratings, they were perplexed about this correlation. Based on teacher comments, it can be concluded that TEP (modified) does not clearly reflect if, or how, teacher attributes impact their perceptions of NCTEP.

Findings from this study support the conclusions that teacher perceptions of the impact of NCTEP on professional growth, teacher effectiveness, and attitudes toward teaching differ based on years of teaching experience and grade level taught at the time of the last evaluation. In all three outcome rating categories, teachers with the most experience reported the lowest impact of NCTEP. Thirty-nine percent of teachers with

20 or more years of experience indicated a low ranking impact to changes in their teaching strategies/behaviors; 51.6% of this group gave a low ranking impact on teacher effectiveness; and 40.3% reported a low impact on their attitudes toward teaching. Teachers with less experience, particularly those with 10 or fewer years, reported a higher impact of NCTEP on their teaching strategies/behaviors (40%); teacher effectiveness (37.7%); and attitudes toward teaching 45.9%. Additionally, data support the conclusion that the grade level taught at the time of last evaluation influences teacher perceptions of NCTEP. Generally, teachers in Grades 9 and above report lower impact ratings on teacher growth (47.3%), teacher effectiveness (53.6%), and attitudes toward teaching (44.6%) than any other group. It can be concluded that teachers who are more experienced and who teach grade levels 9 and above perceive that NCTEP has a lesser impact on their professional growth, teacher effectiveness, and their attitudes toward teaching than is reported by other groups of teachers.

In this survey, information was collected in three methods: (1) through quantitative data generated from TEP (modified); (2) through narrative, optional comments to selected TEP (modified) questions; and (3) from follow-up interviews with voluntary participants. Narrative comments reported on TEP (modified) were overwhelmingly negative in tone; however, almost half (47.2%) of the total 197 survey respondents leaned toward a high quality rating for the overall quality of NCTEP. Interviewee responses also indicated a more positive tone than the online comments, stating that NCTEP encouraged reflection (62.5%) and challenged them to be better teachers (36.4%). Interviewees did, however, express concerns about the potential use of information from NCTEP and the pressure created by the process (36.4%). Based on the inconsistency between the negative tone of comments from the online survey, the overall

quality ratings from the survey, and comments from follow-up interviews, it can be concluded that teacher perceptions of NCTEP are complex. It is possible that the method of data collection influenced how teachers responded. For instance, the assurance of anonymity may have contributed to the negative tone of comments in the online survey as teachers felt comfortable being frank in their responses. The timing of the online survey, which was near the end of the school year, may also have added to the negativity. This conclusion is supported in one interviewee's statement that she was *hostile* when she completed the online survey; but when the follow-up interview was conducted in July, she discussed her experiences with NCTEP in a more positive tone.

From teacher comments on TEP (modified) and in follow-up interviews, it can be concluded that there is a common perception that the evaluation process does not include all aspects that contribute to a teacher's performance. Teachers report that student factors such as motivation, attendance, behavior, and intellectual qualities should be taken into consideration in NCTEP. Teacher factors, including relational qualities, personal characteristics, and work ethic, were other areas that respondents identified as important to consider in the evaluation process. One interviewee offered a suggestion for addressing some of these elements and improving NCTEP by adding "an open section for principals to address criteria that are not explicitly on the evaluation," stating that this would be a way to "inject a little bit more of humanity into the process." This suggestion is supported through RttT recommendations by Davis et al. (2015) to add a standard to NCTEP that addressed the relational aspects of teaching.

Analysis of online survey data and follow-up interviews demonstrated that one attribute of NCTEP that is closely connected to teacher growth is feedback. From correlation data and teacher statements, it can be concluded that when teachers receive

quality feedback, it promotes productive conversations that encourage self-reflection and personal goal setting. PDP and related professional development, which are elements of the context of the evaluation, also contributed to professional growth; but when relevant professional development is not available, teachers report becoming frustrated with the evaluation process. It can also be concluded that the evaluator's role in the evaluation process is critical to whether teachers perceive NCTEP positively or negatively. This was evident in online comments and follow-up interviews when the same attributes were reported to have both positive and negative impacts on growth, depending on teacher experiences with the evaluator. Research by Leggett (2015), who examined the influence of the North Carolina principal characteristics on teacher ratings, recommended that further study be conducted to determine how principals' prior instructional experiences influence their evaluations of teachers. Qualitative data from this study support Leggett's recommendation that more research into the influence of principal characteristics on teacher ratings is important in understanding how to improve the teacher evaluation process.

A significant number of survey and interview responses indicated that teachers perceive that NCTEP has little or no impact on either their effectiveness or their attitudes toward teaching. For those respondents who identified NCTEP as impacting their perceptions of teacher effectiveness and attitudes toward teaching, the attributes of evaluation procedures and context of the evaluation were related to their perceptions; however, there was no consistency in these responses as to whether NCTEP's impact was positive or negative. For example, accountability, peer observations, and rater bias were named as reasons for positive and negative impacts on teacher effectiveness. Furthermore, teachers listed factors other than NCTEP as having more impact on their

teacher effectiveness and attitudes toward teaching. Some of these factors included

collaboration with colleagues, personal study and research, concern for students, and the

desire to improve as a professional. It can be concluded that while teachers in this study

reported that they understood the necessity of NCTEP for accountability, they generally

did not see NCTEP as helpful in improving their effectiveness as a teacher, nor did they

report that NCTEP changed their attitudes toward teaching.

The NCTEP manual states that the intended purpose of NCTEP is "to assess the

teacher's performance in relation to the North Carolina Professional Teaching Standards

and to design a plan for professional growth" (Public Schools of North Carolina, 2015, p.

4). Survey results showed that 34.3% of respondents viewed the intended purpose of

NCTEP as balanced between accountability and growth. Less than a fourth (about 23%)

felt that the role was more toward accountability, and a majority (about 41%) leaned

more toward teacher growth. When asked to discuss these results, interviewee comments

supported the distribution of responses found in TEP (modified). The conclusion can be

drawn that teachers in this study perceive that NCTEP is meeting both accountability and

growth purposes, although they are divided on which purpose is, or should be, the most

prominent.

Recommendations for Further Study

Throughout the RttT grant funding period, periodic status reports were required

by the United States Department of Education to determine the impact and effectiveness

of North Carolina's initiatives. Four targeted areas, designated *pillars*, categorized

educational objectives (NCDPI, 2015). The effectiveness of the evaluation process was

considered under the pillar of Great Principals and Leaders. Marks et al. (2015) reported

that RttT initiatives under this pillar were aimed at enhancing educator effectiveness and

supporting teacher growth by developing teacher capacity to provide data-driven instruction.

Findings presented in the final report on the overall impact of North Carolina's RttT initiatives revealed that higher levels of quality teaching were not related to either principal ratings on Standards 1-5 or to EVAAS ratings (Marks et al., 2015). No statistically significant findings demonstrated increased effectiveness when comparing changes in performance to the rate of improvement prior to RttT (Marks et al., 2015). Additionally, teacher perceptions about the fairness of teacher evaluations decreased between 2012 and 2014 (Marks et al., 2015). Marks et al. (2015) noted important limitations to RttT study, including,

1. Difficulty in establishing a valid comparison group against which to measure initiatives;

2. Difficulty in determining the true impact of RttT initiatives since other services may also have affected outcomes;

3. Difficulty in discerning the impact of changes in leadership and economic conditions on RttT outcomes that were independent of the initiatives themselves; and

4. The limited time frame from "initial implementation to outcome measurement [that] may not yet have been sufficient to capture many of the initiatives' eventual intended effects," stating that: "determining the impact of a multi-year, state-level intervention comprised of multiple initiatives in multiple locations with a variety of implementers is challenging at best" (p. 32).

Given these limitations of RttT reporting, further study of the impact of NCTEP on teacher effectiveness, teacher growth, and attitudes toward teaching was deemed

necessary by this researcher. The potential for continuing investigation of NCTEP's impact is evident in the following recommendations for study.

Recommendation 1. Exploration of the relationship between teacher attributes and the perceived impact of NCTEP on teacher effectiveness, teacher growth, and teachers' attitudes toward teaching should be studied further.

Stiggins and Duke (1988) stated that teacher attributes refer to instructional competence, personal expectations, openness to suggestions, orientation to change, subject knowledge, and experience levels of teachers. Understanding how these characteristics influence teacher perceptions of NCTEP should be studied further as data from this study did not demonstrate a relationship between teacher attributes and the outcome ratings; however, interviewee comments showed that this lack of relationship was puzzling to them. Interviewees gave diverse reasons why there was no qualitative relationship. They speculated that teacher self-ratings were incorrect, the design of the survey contributed to this data, teacher attitudes about the instrument or the evaluator influenced the outcome, or NCTEP does not clearly address teacher attributes.

Recommendation 2. Based on survey comments and follow-up interviews, the same attributes of the evaluation were noted to have both positive and negative impacts on professional growth. Identification of the causal factors for these conflicting impacts is an area for further study.

When examining study responses, this researcher noted that teachers' unique experiences with NCTEP influenced their opinions about whether the same attributes were perceived positively or negatively. For example, the role of the evaluator, the context of the evaluation, and evaluation procedures were described both positively and negatively by various respondents. Inconsistencies in the tone of responses were also

evidenced between the online survey comments, qualitative survey data, and follow-up interviews. It is recommended that case studies be conducted to explore the experiences of individual teachers with NCTEP. This may reveal why the same attributes produced contrasting viewpoints from teachers and identified policy implications that will enhance NCTEP's potential to increase teacher growth and effectiveness as well as positively impact teacher attitudes toward teaching.

Recommendation 3. Findings from this study support the conclusions that teacher perceptions of the impact of NCTEP on professional growth, teacher effectiveness, and attitudes differ based on years of teaching experience and grade level taught at the time of the last summative evaluation. Additional investigation into why teacher perceptions differ according to these demographics is suggested.

Overall, this study found that teachers with more experience were less likely to report higher impacts on teacher effectiveness and attitudes toward teaching than other groups; and teachers with 10 years or less experience reported higher impacts than other groups. Teachers in Grades 9 and above reported lesser impacts on attitudes toward teaching than any other group, while Grades 6-8 teachers reported the highest impact ratings for attitude than other group. Further study of these results may reveal commonalities of the evaluation experiences of these groups that influence the impact of NCTEP.

Recommendation 4. Study responses indicated that teachers viewed accountability and growth as dual purposes of NCTEP. Additionally, interviewees commented that they felt that evaluation for accountability had the potential to be beneficial; however, they stated that the process was only useful when it was pertinent to their work in the classroom with students. More research is needed to determine if, and

how, NCTEP is relevant to the actual work of the teacher in the classroom.

Generally, teacher responses demonstrated that NCTEP was viewed as having either little to no impact on their teaching effectiveness and attitudes toward teaching or NCTEP was viewed negatively. Comments from survey respondents described the evaluation process as "busy work" and a "bureaucratic hurdle to cross." Analysis of quantitative data indicated that a relationship existed between the context of the evaluation and teacher perceptions of NCTEP and was investigated in follow-up interviews. Follow-up interviews revealed that teachers were divided as to whether the primary purpose of NCTEP should be growth or accountability. Concerns about how information from NCTEP would be used and increased stress levels as a result of the evaluation process were also noted. From online survey comments, it was determined that teachers felt that consideration of factors outside of their control, such as student motivation and attendance, should be considered during the evaluation process. These data show that teacher confidence in the usefulness of the evaluation process is questionable. This opinion is summed up in one teacher's statement which noted that NCTEP has the potential of improving teacher effectiveness but that when the process is put into practice, it is not achieving its potential. Further study of how NCTEP can improve its potential and increase teacher growth and teacher effectiveness is warranted.

Summary

The findings of this study suggest that perspectives impact perceptions. Teachers' unique experiences with NCTEP influence their perceptions of the impact NCTEP has on teacher effectiveness, professional growth, and their attitudes toward teaching. Teachers with more experience report lower impact ratings on teacher effectiveness and attitudes toward teaching than those with lesser experience, and teachers in higher grades (9 and

174

above) tend to follow this trend. Professional growth is related to the key evaluation attribute of feedback; however, the impact of this attribute can be complicated when teachers are distrusting of the evaluator's ability to provide productive feedback or when professional development is unavailable to support their growth as a practitioner.

Teachers in this study recognized that accountability and growth are two purposes for teacher evaluation, and they reported that NCTEP reflects both purposes. Teacher responses are consistent with the views of evaluation researchers (Darling-Hammond et al., 1983; Flowers et al., 2000; Marzano, 2012; Weisberg et al., 2009) in indicating that fulfilling both purposes through one system is complex. This study revealed opportunities for continued research into how educational policies, particularly those involving the attributes of the evaluator, evaluation procedures, and the context of the evaluation, can be improved to enhance NCTEP's potential to increase teacher growth and effectiveness as well as positively impact teacher attitudes toward teaching.

References

American Recovery and Reinvestment Act of 2009, Pub. L. 111-5, § 14005-6, Title XIV.

Balls, J. D., Eury, A. D., & King, J. C. (2011). *Rethink, rebuild, rebound: A framework for shared responsibility and accountability in education* (2d ed.). Boston, MA: Pearson Learning Solutions.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall, Inc.

Beckham, J. C. (1981). *Legal aspects of teacher evaluation.* Topeka, KS: National Organization on Legal Problems of Education.

Betebenner, D., Braun, H., Corcoran, S., Darling-Hammond, L., Friedman, J., & Goldhaber, D. (2012). *Learning from recent advances in measuring teacher effectiveness.* Washington, DC: Institute of Education Sciences.

Bradshaw, L. K. (1996). Alternative teacher performance appraisal in North Carolina: Developing guidelines. Retrieved from http://eric.ed.gov/?id=ED400255

Bradshaw, L. K. (2002). Local district implementation of state mandated teacher evaluation policies and procedures: The North Carolina case. *Journal of Personnel Evaluation in Education, 16*(2), 113-127.

Brain, G. B. (1978). Political issues in teacher education and certification. In W. I. Israel (Ed.), *Political issues in education: A report of the chief state school officers summer institute* (pp. 117-125). Washington, DC: Chief Council of State School Officers.

Brickell, H. M. (1978). Political issues in competency testing. In W. I. Israel (Ed.), *Political issues in education: A report of the chief state school officers summer institute* (pp. 65-81). Washington, DC: Chief Council of State School Officers.

Bridges, E. M., & Groves, B. R. (1999). The macro- and micropolitics of personnel evaluation: A framework. *Journal of Personnel Evaluation in Education, 13*(4), 321.

Campbell, J. M. (1978). Political issues in the relation between education and state government. In W. I. Israel (Ed.), *Political issues in education: A report of the 1978 Chief State School Officers Summer Institute* (pp. 33-38). Washington, DC: Chief Council of State School Officers.

Collins, C. (2014). Houston, we have a problem: Teachers find no value in the SAS education value-added assessment system (EVAAS®). *Education Policy Analysis Archives, 22*(98).

Corn, J., Bryant, L., Maxfield, J., Walton, M., Patel, R., Weiss, S. P., . . . Smart, A. (2013). *Executive summary of second annual race to the top evaluation report: Part I - statewide face-to-face professional development formative evaluation.* Retrieved from http://cerenc.org/wp-content/uploads/2011/11/ES-2nd-Annual-PD-Report_statewide_02-01-13.pdf

Corn, J., Halstead, E., Kellogg, S., Kleiman, G., Tagswold, J., Townsend, M., . . . Lewis, K. (2012). *Building LEA and regional professional development capacity: First annual evaluation report.* Retrieved from http://cerenc.org/wp-content/uploads/2012/01/PD_1st_Year_Report_1-5-121.pdf

Corn, J., Smart, A., Bryant, L., Maxfield, J., Walton, M., Patel, R., . . . Kellogg, S. (2013). *Second annual race to the top professional develoopment evaluation report: Part II - local outcomes baseline study.* Retrieved from http://cerenc.org/wp-content/uploads/2011/11/2nd-Annual-PD-Report_local-outcomes_02-01-13.pdf

Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice.* Alexandria, VA: Assoc. for Supervision and Curriculum Development.

Darling-Hammond, L. (2013). *Getting teacher evaluation right: What really matters for effectiveness and improvement.* New York, NY: Teachers College Press.

Darling-Hammond, L. (2014). One piece of the whole: Teacher evaluation as part of a comprehensive system for teaching and learning. *American Educator, Spring 2014*, 13, 44.

Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *The Phi Delta Kappan, 93*(6), 8-15. doi:10.1177/003172171209300603

Darling-Hammond, L., Wise, A. E., & Pease, S. R. (1983). *Teacher evaluation in the organizational context: A review of the literature.* Santa Monica, CA: Rand Corporation.

Databases on State Teacher and Principal Evaluation Policies. (n.d.). *Center on great teachers & leaders at American Institutes for Research.* Retrieved from http://resource.tqsource.org/stateevaldb/

Davey, B. (1991). Evaluating teacher competence through the use of performance assessment tasks: An overview. *Journal of Personnel Evaluation in Education, 5*(2), 121-32.

Davis, C. R., Bangert, L., Comperatore, A. N., & Smalenberger, M. (2015). Teacher and principal perceptions of the North Carolina educator evaluation system: Final report. *Consortium for Educational Research and Evaluation - North Carolina (CERENC).* Retrieved from http://cerenc.org/rttt-evaluation/teacher-and-principal-effectiveness/

Dewey, J. (2014). The democratic faith and education. *The Antioch Review,* 783-794. doi:10.7723/antiochreview.72.4.0783

Donaldson, M. L. (2012). *Teachers' perspectives on evaluation reform.* Washington, DC: Center for American Progress. Retreived from https://cdn.americanprogress.org/wp-content/uploads/2012/12/TeacherPerspectives.pdf

Doyle, D., & Han, J. G. (2012). *Measuring teacher effectiveness: A look "under the hood" of teacher evaluation in 10 sites.* New Haven, CT: ConnCAN.

Easton, L. B. (2012). Guide to evaluating teacher evaluation. *Supplement to Phi Delta Kappan: Professional Development Discussion Guide, 6*(93), 5.

Eller, J. F., & Eller, S. A. (2015). *Score to soar: Moving teachers from evaluation to professional growth*. Bloomington, IN: Solution Tree Press.

Flowers, C., Testerman, J., Hancock, D., & Algozzine, B. (2000). *Technical manual: NC experienced teacher summative evaluation system: TPAI-2000.* Charlotte, NC: Department of Educational Administration, Research, and Technology.

Frechtling, J., & Sharp, L. (1997). *User-friendly handbook for mixed method evaluations.* Arlington, VA: National Science Foundation.

Gabriel, R., & Allington, R. (2012). The MET project: The wrong 45 million dollar question. *Educational Leadership, 70*(3), 44-49.

Goldhammer, R. (1969). *Clinical supervision: Special methods for the supervision of teachers.* Oxford, England: Holt, Rinehart, & Winston.

Griffin, L. (2013). Charlotte Danielson on teacher evaluation and quality. *School Administrator, 70*(1), 27-31.

Gullickson, A. R., & Howard, B. (2009). *The personnel evaluation standards: How to assess systems for evaluating educators*. Thousand Oaks, CA: Corwin.

Haertel, E. (1986). *The valid use of student performance measures for teacher evaluation.* Washington, DC: American Educational Research Association.

Hallinger, P., Heck, R. H., & Murphy, J. (2014). Teacher evaluation and school improvement: An analysis of the evidence. *Educational Assessment, Evaluation and Accountability, 26*(1), 5-28. doi://dx.doi.org.ezproxy.gardner-webb.edu/10.1007/s11092-013-9179-5

Halperin, S. (1978). Political issues in establishing education priorities. In W. I. Israel (Ed.), *Political issues in education: A report of the state chief school officers summer institute* (pp. 7-22). Washington, DC: Chief Council of State School Officers.

Harris, D. N., Ingle, W. K., & Rutledge, S. A. (2014). How teacher evaluation methods matter for accountability: A comparative analysis of teacher effectiveness ratings by principals and teacher value-added measures. *American Educational Research Journal, 51*(1), 73-112.

Hechinger, F. M. (1978). Political issues in education: Reflections and directions. In W. I. Israel (Ed.), *Political issues in education: A report of the chief state school officers summer institute* (pp. 127-135). Washington, DC: Chief Council of State School Officers.

Henry, G. T., Campbell, S. L., Thompson, C. L., & Townsend, L. W. (2014). *Evaluation of district and school transformation school-level coaching and professional development activities.* North Carolina: Consortium for Educational Research and Evaluation - North Carolina. Retrieved from https://publicpolicy.unc.edu/files/2015/12/Evaluation-of-District-and-School-Transformation-School-Level-Coaching-and-Professional-Development-Activities-October-2014.pdf

Henry, G. T., & Guthrie, E. (2015). *An evaluation of the North Carolina educator evaluation system and the student achievement growth standard: 2010-11 through 2013-14.* North Carolina: Consortium for Educational Research and Evaluation - North Carolina. Retrieved from http://cerenc.org/wp-content/uploads/2015/09/0-FINAL-Evaluation-of-NC-Teacher-Evaluation-9-3-15.pdf

Hirsh, S. (2015). New bill offers a good start on defining PD. Retrieved from http://learningforward.org/publications/blog/learning-forward-blog/2015/12/18/new-bill-offers-a-good-start-on-defining-pd#.VwqZd6QrLIV

Holdzkom, D. (1987). Appraising teacher performance in North Carolina. *Educational Leadership, 44*(7), 40-44.

Hull, J. (2013). Trends in teacher evaluation: How states are measuring teacher performance. *National School Boards Association: Center for Public Education,* 38. Retrieved from http://www.centerforpubliceducation.org/Main-Menu/Evaluating-performance/Trends-in-Teacher-Evaluation-At-A-Glance/Trends-in-Teacher-Evaluation-Full-Report-PDF.pdf

Inman, W. C. (1982a). *Reliability of the teacher and principal personnel performance appraisal instruments.* Raleigh, NC: North Carolina Board of Education.

Inman, W.C. (1982b). *Reactions of teachers, principals, and superintendents to 1981-82 performance appraisals.* Raleigh, NC: North Carolina Board of Education.

Interstate New Teacher Assessment and Support Consortium. (1992). *Model standards for beginning teacher licensing and development: A resource for state dialogue. draft for comments.* Washington, DC: Interstate New Teacher Assessment and Support Consortium.

Israel, W. I. (1978). *Political issues in education: A report of the 1978 Chief State School Officers Summer Institute.* Washington, DC: Council of Chief State School Officers.

Joint Committee on Standards for Educational Evaluation. (1988). *The personnel evaluation standards*. Newbury Park, CA: Sage.

Katz, D. S. (2016). Growth models and teacher evaluation: What teachers need to know and do. *Kappa Delta Pi Record, 52*(1), 11. doi:10.1080/00228958.2016.1123039

Lee, L. (2011). What did the teachers think? Teachers' responses to the use of value-added modeling as a tool for evaluating teacher effectiveness. *Journal of Urban Learning, Teaching, and Research, 7*, 97-103.

Leggett, S. (2015). *The influence of principal and school characteristics on principal ratings of teachers using the north carolina teacher evaluation instrument.* Retrieved from http:// search.proquest.com/docview/1648413303

Lewis, A. C. (1982). *Evaluating educational personnel*. Arlington, VA: American Association of School Administrators.

Lofton, G. G., Hill, F., & Claudet, J. G. (1997). Can state-mandated teacher evaluation fulfill the promise of school improvement? Events in the life of one school. *Journal of Personnel Evaluation in Education, 11*(2), 127-137. doi:1007937303372

Lynn, H., Barrett, N., Marks, J., Comperatore, A., Henry, G. T., & Guthrie, J. (2013). *Executive summary of measures of student growth in the North Carolina educator evaluation system: Formative evaluation report.* North Carolina: Consortium for Educational Research and Evaluation - North Carolina. Retrieved from http://cerenc.org/wp-content/uploads/2011/10/ES-Formative-report-TLEE-measures-of-student-growth-9-x-13.pdf

Lyttle, L. (2011). *Has the no child left behind law produced more qualified teachers?* Online Submission. Retrieved from http://eric.ed.gov/?id=ED536289

Marks, J. T., Fuller, S. C., Guthrie, J. E., Henry, G. T., & Stallings, D. T. (2015). *North Carolina race to the top: Overall impact and implementation findings - final report.* Retrieved from http://cerenc.org/wp-content/uploads/2015/09/ES-FINAL-Overall-9-3-15.pdf

Marzano, R. J. (2012). The two purposes of teacher evaluation. *Educational Leadership, 70*(3), 14-19. Retrieved from http://www.ascd.org/publications/educational-leadership/nov12/vol70/num03/The-Two-Purposes-of-Teacher-Evaluation.aspx

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

McLaughlin, M. W. (1982). *A preliminary investigation of teacher evaluation practices.* Santa Monica, CA: The Rand Corporation.

McNergney, R. F., & Imig, S. R. (2003). Teacher evaluation: Overview. In J. W. Guthrie (Ed.), *Encyclopedia of education* (2nd ed., pp. 2453-2457). NY: Macmillan.

Mead, S. (2012). *Recent state action on teacher effectiveness: What's in state laws and regulations?* Retrieved from http://bellwethereducation.org/sites/default/files/legacy/2012/08/RSA-Teacher-Effectiveness.pdf

Medley, D. M. (1982). *Teacher competency testing and the teacher educator*. (s.l.) Charlottesville, VA: Bureau of Educational Research, School of Education, University of Virginia.

MET Project. (2013). *Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET project's three-year study - policy and practice brief.* Seattle, WA: Bill & Melinda Gates Foundation.

National Center for Educational Evaluation and Regional Assistance. (2014). *State requirements for teacher evalution policies promoted by race to the top.* Retrieved from https://ies.ed.gov/ncee/pubs/20144016/pdf/20144016.pdf

NCLEG. (2010). *North carolina's race to the top plan: Fiscal brief.* Retrieved from http://www.ncleg.net/fiscalresearch/fiscal_briefs/Fiscal_Briefs_PDFs/Race%20to%20the%20Top%20Fiscal%20Brief_FINAL.pdf

No child left behind act. (2001). 20 U.S.C.§ 6301. Retrieved from http://www.2.ed.gov/policy/elsec/leg/esea02/107-110.pdf

North Carolina Department of Public Instruction. (2015). *North Carolina race to the top closeout report: Executive summary.* Retrieved from https://www2.ed.gov/programs/racetothetop/state-reported-sharing/ncexsumm.pdf

North Carolina State Board of Education. (n.d.) *Chapter six: The charter schools act & the excellent schools act.* North Carolina State Board of Education. Retrieved from http://stateboard.ncpublicschools.gov/about-sbe/history/chapter-six

O'Hanlon, J., & Mortenson, L. (1977). Approaches to teacher evalaution. *CEDR Quarterly, 10*(4), 3-7.

Orphanos, S. (2014). What matters to principals when they evaluate teachers? Evidence from cyprus. *Educational Management Administration & Leadership, 42*(2), 243-258. doi:10.1177/1741143213499262

Pathe, S., & Choe, J. (2013). A brief overview of teacher evaluation controversies. Retrieved from http://www.pbs.org/newshour/rundown/teacher-evaluation-controversies/

Public Schools of North Carolina. (n.d.a). Educator effectiveness model. Retrieved from http://www.ncpublicschools.org/effectiveness-model/ncees/

Public Schools of North Carolina. (n.d.b). *Future-ready schools: Preparing students for the 21st century - biennial report (2004-2006).* Raleigh, NC: Retrieved from http://stateboard.ncpublicschools.gov/resources/biennial-reports/2004-06biennial-report.pdf

Public Schools of North Carolina. (n.d.c). North Carolina educator evaluation system (NCEES) wiki. Retrieved from http://ncees.ncdpi.wikispaces.net/NCEES+Wiki

Public Schools of North Carolina. (n.d.d). North Carolina's analysis of student work wiki. Retrieved from http://ncasw.ncdpi.wikispaces.net/

Public Schools of North Carolina. (n.d.e). *Student growth.* Educator Effectiveness Model. Retrieved from http://www.dpi.state.nc.us/effectiveness-model/student-growth/

Public Schools of North Carolina. (2009). *North Carolina teacher evaluation process.* Denver, CO: McREL.

Public Schools of North Carolina. (2010). *North Carolina RTTT.* Retrieved from http://www.dpi.state.nc.us/rttt/

Public Schools of North Carolina. (2012a). *The ABCs of public education: 2012 accountability report background packet.* Retrieved from http://www.ncpublicschools.org/docs/accountability/reporting/abc/2011-12/backgroundpacket.pdf

Public Schools of North Carolina. (2012b). *North Carolina teacher evaluation process, 2012.* Denver, CO: McREL.

Public Schools of North Carolina. (2013a). *Measuring student learning for educator effectiveness: A guide to the use of student growth data in the evaluation of North Carolina teachers.* Retrieved from http://www.dpi.state.nc.us/docs/effectiveness-model/student-growth/measuring-growth-guide.pdf

Public Schools of North Carolina. (2013b). *North Carolina professional teaching standards.* Retrieved from http://www.ncpublicschools.org/docs/effectiveness-model/ncees/standards/prof-teach-standards.pdf

Public Schools of North Carolina. (2014). They just want to fire me! NC's educator effectiveness myths debunked. Retrieved from http://www.ncpublicschools.org/docs/effectiveness-model/resources/effectiveness-myths.pdf

Public Schools of North Carolina. (2015). *North Carolina teacher evaluation process, 2015.* Retrieved from http://www.ncpublicschools.org/docs/effectiveness-model/ncees/instruments/teach-eval-manual.pdf

Rooney, P., Hussar, W., Planty, M., Choy, S., Hampden-Thompson, G., & Provasnik, S. (2006). The condition of education. National Center for Education Statistics: Washington, DC, NCES 2006-071, 409. Retrieved from http://nces.ed.gov/pubs2006/2006071.pdf

Rose, R. A., Henry, G. T., & Lauen, D. L. (2012). *Comparing value-added models for estimating individual teacher effects on a statewide basis: Executive Summary.* Retrieved from http://cerenc.org/wp-content/uploads/2011/10/ES_VAM_Technical-Report_8-27-12.pdf

Sanders, W. L. (2000). Value-added assessment from student achievement data: Opportunities and hurdles. *Journal of Personnel Evaluation in Education, 14*(4), 329.

Sanders, W. L., & Horn, S. P. (1998). Research findings from the tennessee value-added assessment system (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education, 12*(3), 247-256. doi:1008067210518

SAS. (2015). *Beyond a single estimate of effectiveness: Conquering education data complexity in reporting.* Retrieved from http://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/beyond-single-estimate-effectiveness-107722.pdf

Sawchuk, S. (2015). Teacher evaluation heads to the courts. *Education Week, 7*(35), 1-6. Retrieved from http://www.edweek.org/ew/section/multimedia/teacher-evaluation-heads-to-the-courts.html?cmp=eml-enl-eu-news2

Sclan, E., & Darling-Hammond, L. (1992). *Beginning teacher performance evaluation: An overview of state policies. Trends and issues paper no. 7.* Washington, DC: ERIC Clearinghouse on Public Education.

Shakman, K., Riordan, J., Sanchez, M. T., Cook, K. D., Fournier, R., & Brett, J. (2012). *An examination of performance-based teacher evaluation systems in five states: Issues & answers. REL 2012-no. 129.* Regional Educational Laboratory Northeast & Islands.

Smart, A., Bryant, L., Kellogg, S., Corn, J., Weiss, S. P., Parker, B., . . . Davis, C. (2015). *Fourth annual race to the top professional development evaluation report.* Retrieved from http://cerenc.org/wp-content/uploads/2015/03/ES-FINAL-Final-Annual-RttT-PD-Report-3-5-15.pdf

Stacey, D. C., Holdzkom, D., & Kuligowski, B. (1989). Effectiveness of the North Carolina teacher performance appraisal system. *Journal of Personnel Evaluation in Education, 3*(1), 79-106. doi:10.1007/BF00125089

Stiggins, R. (1986). Teacher evaluation: Accountability and growth systems-different purposes. *NASSP Bulletin, 70*(490), 51-58. doi:10.1177/019263658607049015

Stiggins, R. J., & Duke, D. L. (1988). *The case for commitment to teacher growth: Research on teacher evaluation*. Albany, NY: State University of New York Press.

Stronge, J. H., & Tucker, P. D. (1999). The politics of teacher evaluation: A case study of new system design and implementation. *Journal of Personnel Evaluation in Education, 13*, 339-360.

Taylor, E. S., & Tyler, J. H. (2012). Can teacher evaluation improve teaching? *Education Next, 12*(4), 78-84. Retrieved from http://educationnext.org/can-teacher-evaluation-improve-teaching/

The Teaching Channel (2013). *Multiple measures of teaching effectiveness.* Retrieved from https://teachingchannel.org/videos/improving-teacher-performance

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

Vagias, W. M. (2006). *Likert-type scale response anchors.* Clemson, SC: Clemson University.

Weick, K. E. (1982). Administering education in loosely-coupled systems. *Phi Delta Kappan, 63*(10), 673-676.

Weisberg, D., Sexton, S., Mulhern, J., Keeling, D., Schunck, J., Palcisco, A., & Morgan, K. (2009), The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. Brooklyn, NY: The New Teacher Project. Retrieved from http://widgeteffect.org/downloads/TheWidgetEffect.pdf

Wetzel, W. A. (1929). Scientific supervision and curriculum-building. *The School Review 37*(2), 117-123. doi: DOI: 10.1086/438807

Williams, J. (2009). *McREL's teacher evaluation*. Denver, CO: Mid-continent Research for Education and Learning.

Wise, A. E., Darling-Hammond, L., McLaughlin, M. W., & Bernstein, H. T. (1984). *Teacher evaluation: A study of effective practices*. Santa Monica, CA: Rand.

Appendix A

Teacher Performance Appraisal Instrument (TPAI)

Criteria of the North Carolina Teacher Performance Appraisal Instrument
(Adapted form Stacey et al., 1989, p. 104-106)

1. Major Function: *Management of Instructional Time*

> 1.1 Teacher has materials, supplies, and equipment ready at the start of the lesson or instructional activity.
> 1.2 Teacher gets the class started quickly.
> 1.3 Teacher gets students on task quickly at the beginning of each lesson or instructional activity.
> 1.4 Teacher maintains a high level of student time-on-task.

2. Major Function*: Management of Student Behavior*

> 2.1 Teacher has established a set of rules and procedures that govern the handling of routine administrative matters.
> 2.2 Teacher has established a set of rules and procedures that govern student verbal participation and talk during different types of activities--whole- class instruction, small-group instruction, and so on.
> 2.3 Teacher has established a set of rules and procedures that govern student movement in the classroom during different types of instructional activities.
> 2.4 Teacher frequently monitors the behavior of all students during whole- class, small-group, and seatwork activities and during transitions between instructional activities.
> 2.5 Teacher stops inappropriate behavior promptly and consistently, yet maintains the dignity of the student.

3. Major Function: *Instructional Presentation*

> 3.1 Teacher begins lesson or instructional activity with a review of previous material.
> 3.2 Teacher introduces the lesson or instructional activity and specifices learning objectives when appropriate.
> 3.3 Teacher speaks fluently and precisely.
> 3.4 Teacher presents the lesson or instructional activity using concepts and language understandable to the students.
> 3.5 Teacher provides relevant examples and demonstrations to illustrate concepts and skills.
> 3.6 Teacher assigns tasks that students handle with a high rate of success.
> 3.7 Teacher asks appropriate levels of questions that students handle with a high rate of success.
> 3.8 Teacher conducts lesson or instructional activity at a brisk pace, slowing presentations when necessary for student understanding but avoiding unnecessary slowdowns.
> 3.9 Teacher makes transitions between lessons and between instructional activities within lessons efficiently and smoothly.
> 3.10 Teacher makes sure that the assignment is clear.

3.11 Teacher summarizes the main point(s) of the lesson at the end of the lesson or instructional activity.

4. Major Function: *Instructional Monitoring of Student Performance*

4.1 Teacher maintains clear, firm, and reasonable work standards and due dates.
4.2 Teacher circulates during classwork to check all students' performance.
4.3 Teacher routinely uses oral, written, and other work products to check student progress.
4.4 Teacher poses questions clearly and one at a time.

5. Major Function: *Instructional Feedback*

5.1 Teacher provides feedback on the correctness or incorrectness of in-class work to encourage student growth.
5.2 Teacher regularly provides prompt feedback on assigned out-of-class work.
5.3 Teacher affirms a correct oral response appropriately, and moves on.
5.4 Teacher provides sustaining feedback after an incorrect response or no response by probing, repeating the question, giving a clue, or allowing more time.

6. Major Function: *Facilitating Instruction*

6.1 Teacher has an instructional plan that is compatible with the school and system-wide curricular goals.
6.2 Teacher uses diagnostic information obtained from tests and other assessment procedures to develop and revise objectives and/or tasks.
6.3 Teacher maintains accurate records to document student performance.
6.4 Teacher has instructional plan that matches/aligns objectives, learning strategies, assessment, and student needs at the appropriate level of difficulty.
6.5 Teacher uses available human and material resources to support the instructional program.

7. Major Function: *Communicating Within the Educational Environment*

7.1 Teacher treats all students in a fair and equitable manner.
7.2 Teacher interacts effectively with students, co-workers, parents, and community.

8. Major Function: *Performing Non-Instructional Duties*

8.1 Teacher carries out non-instructional duties as assigned and/or as need is perceived.
8.2 Teacher adheres to established laws, policies, rules, and regulations.
8.3 Teacher follows a plan for professional development and demonstrates evidence of growth.

Rating Scale of the North Carolina Teacher Performance Appraisal Instrument

Level of Performance

6. *Superior*
Performance within this function area is consistently outstanding. Teaching practices are demonstrated at the highest level of performance. Teacher continuously seeks to expand scope of competencies and constantly undertakes additional, appropriate responsibilities.

5. *Well Above Standard*
Performance within this function areas is frequently outstanding. Some teaching practices are demonstrated at the highest level, while others are at a consistently high level. Teacher frequently seeks to expand scope of competencies and often undertakes additional, appropriate responsibilities.

4. *Above Standard*
Performance within this function area is frequently high. Some teaching practices are demonstrated at a high level, while others are at a consistently adequate/acceptable level. Teacher sometimes seeks to expand scope of competencies and occasionally undertakes additional, appropriate responsibilities.

3. *At Standard*
Performance within this function area is consistently adequate/acceptable. Teaching fully meet all performance expectations at an acceptable level. Teacher maintains an adequate scope of competencies and performs additional responsibilities as assigned.

2. *Below Standard* Performance within this function area is sometimes inadequate/unacceptable and needs improvement. Teacher requires supervision and assistance to maintain an adequate scope of competencies, and sometimes fails to perform additional responsibilities as assigned.

1. *Unsatisfactory*
Performance within this function area is consistently inadequate/unacceptable, and most practices require considerable improvement to fully meet minimum performance expectations. Teacher requires close and frequent supervision in the performance of all responsibilities.

Appendix B

TPAI-2000

TPAI-2000 Indicators that have been modified from the original TPAI are in bold. (Flowers et al., 2000, p. 5 – 8)

1. **Management of Instructional Time**

    1.1 Teacher has materials, supplies, and equipment ready at the start of the lesson or instructional activity.
    1.2 Teacher gets the class started quickly.
    1.3 Teacher uses available time for learning and keeps students on task.

2. **Management of Student Behavior**

    2.1 Teacher has established a set of rules and procedures that govern the handling of routine administrative matters.
    2.2 Teacher has established a set of rules and procedures that govern student verbal participation and talk during different types of activities—whole class instruction, small group instruction, etc.
    2.3 Teacher has established a set of rules and procedures that govern student movement in the classroom during different types of instructional activities.
    2.4 Teacher frequently monitors the behavior of all students during whole-class, small group, and seat work activities and during transitions between instructional activities.
    2.5 Teacher stops inappropriate behavior promptly and consistently, yet maintains the dignity of the student.
    2.6 Teacher analyzes the classroom environment and makes adjustments to support learning and enhance social relationships.

3. **Instructional Presentation**

    3.1 Teacher links instructional activities to prior learning.
    3.2 Teacher understands the central concepts, tools of inquiry, and structures of the discipline(s) he or she teaches and creates learning activities that make these aspects of subject matter understandable and meaningful for students.
    3.3 Teacher speaks fluently and precisely.
    3.4 Teacher provides relevant examples and demonstrations to illustrate concepts and skills.
    3.5 Teacher assigns tasks and **asks appropriate levels of questions** that students handle with a high rate of success.
    3.6 Teacher conducts the lesson or instructional activity at a brisk pace, slowing presentations when necessary for student understanding but avoiding unnecessary slowdowns.
    3.7 Teacher makes transitions between lessons and between instructional activities within lessons effectively and smoothly.
    3.8 Teacher makes sure that assignment is clear.
    3.9 The teacher creates instructional opportunities that are adapted to diverse learners.
    3.10 The teacher uses instructional strategies that encourage the development of critical thinking, problem solving, and performance skills.

3.11 The teacher uses technology to support instruction.
3.12 The teacher encourages students to be engaged in and responsible for their own learning.

4. Instructional Monitoring of Student Performance

4.1 Teacher maintains clear, firm, and reasonable work standards and due dates.
4.2 Teacher circulates to check all students' performance.
4.3 Teacher routinely uses oral, written, and other work products to evaluate the effects of instructional activities and to check student progress.
4.4 Teacher poses questions clearly and one at a time.
4.5 Teacher uses student responses to adjust teaching as necessary.

5. Instructional Feedback

5.1 Teacher provides feedback on the correctness or incorrectness of in-class work to encourage student growth.
5.2 Teacher regularly provides prompt feedback on out-of-class work.
5.3 Teacher affirms a correct oral response appropriately and moves on.
5.4 Teacher provides sustaining feedback after an incorrect response by probing, repeating the question, giving a clue, or allowing more time.
5.5 The teacher uses knowledge of effective verbal and non-verbal communication techniques to foster active inquiry, collaboration, and supportive interaction in the classroom.

6. Facilitating Instruction

6.1 Teacher has long- and short-term instructional plans that are compatible with school and district curricular goals, the school improvement plan, the NC Standard Course of Study, and the diverse needs of students and the community.
6.2 Teacher uses diagnostic information obtained from tests and **other formal and informal assessment procedures to evaluate and ensure the continuous intellectual, social, and physical development of the learner.**
6.3 Teacher maintains accurate records to document student performance.
6.4 Teacher understands how students learn and develop and plans appropriate instructional activities for diverse student needs and different levels of difficulty.
6.5 Teacher uses available human and material resources to support the instructional program.

7. Communicating within the Educational Environment

7.1 Teacher treats all students in a fair and equitable manner.
7.2 Teacher participates in the development of a broad vision for the school.
7.3 Teacher fosters relationships with school colleagues, parents, and community agencies to support students' learning and well-being.

8. Performing Non-Instructional Duties

8.1 Teacher carries out non-instructional duties as assigned and/or as need is perceived **to ensure student safety outside the classroom.**
8.2 Teacher adheres to established laws, policies, rules, and regulations.
8.3 Teacher follows a plan for professional development and **actively seeks out opportunities to grow professionally**.
8.4 Teacher is a reflective practitioner who continually evaluates the effects of his or her decisions and actions on students, parents, and other professionals in the learning community.

TPAI Rating Scale

4 Above Standard

Performance is consistently high. Teaching practices are demonstrated at a high level. Teacher seeks to expand scope of competencies and undertakes additional appropriate responsibilities.

3 At Standard

Performance within this function area is consistently adequate/acceptable. Teaching practices fully meet all performance expectations at an acceptable level. Teacher maintains an adequate scope of competencies and performs additional responsibilities as assigned.

2 Below Standard

Performance within this function area is sometimes inadequate/unacceptable and needs improvement. Teacher requires supervision and assistance to maintain an adequate scope of competencies and sometimes fails to perform additional responsibilities as assigned.

1 Unsatisfactory

Performance within this function area is consistently inadequate or unacceptable and most practices require considerable improvement to fully meet minimum expectations. Teacher requires close and frequent supervision in the performance of all responsibilities.

Appendix C

North Carolina State Board of Education
Future Ready Students
Goals for the 21st Century

Future-Ready Students
Goals for the 21st Century
(Adapted from Public Schools of North Carolina, b, p. 17)

The guiding mission of the North Carolina State Board of Education is that every public school student will graduate from high school, globally competitive for work and postsecondary education and prepared for life in the 21st century.

NC Public Schools Will Produce Globally Competitive Students.

• Every student excels in rigorous and relevant core curriculum that reflects what students need to know and demonstrate in a global 21st Century environment, including a mastery of languages, an appreciation of the arts and competencies in the use of technology.

• Every student's achievement is measured with an assessment system that informs instruction and evaluates knowledge, skills, performance and dispositions needed in the 21st Century.

• Every student will be enrolled in a course of study designed to prepare them to stay ahead of international competition.

• Every student uses technology to access and demonstrate new knowledge and skills that will be needed as a life-long learner to be competitive in a constantly changing international environment.

• Every student has the opportunity to graduate from high school with an Associate's Degree or college transfer credit.

NC Public Schools Will Be Led By 21st Century Professionals.

• Every teacher will have the skills to deliver 21st Century content in a 21st Century context with 21st Century tools and technology that guarantees student learning.

• Every teacher and administrator will use a 21st Century assessment system to inform instruction and measure 21st Century knowledge, skills, performance and dispositions.

• Every education professional will receive preparation in the interconnectedness of the world with knowledge and skills, including language study.

• Every education professional will have 21st Century preparation and access to ongoing, high quality professional development aligned with State Boardof Education priorities.

• Every educational professional uses data to inform decisions.

NC Public School Students Will Be Healthy and Responsible.

• Every learning environment will be inviting, respectful, supportive, inclusive and flexible for student success.

• Every school provides an environment in which each child has positive, nurturing relationships with caring adults.

• Every school promotes a healthy, active lifestyle where students are encouraged to make responsible choices.

• Every school focuses on developing strong student character, personal responsibility and community/world involvement.

• Every school reflects a culture of learning that empowers and prepares students to be life-long learners.

Leadership Will Guide Innovation in NC Public Schools.

• School professionals will collaborate with national and international partners to discover innovative transformational strategies that will facilitate change, remove barriers for 21st Century learning and understand global connections.

• School leaders will create a culture that embraces change and promotes dynamic, continuous improvement.

• Educational professionals will make decisions in collaboration with parents, students, businesses, education institutions, and faith-based and other community and civic organizations to impact student success.

• Public school professionals will collaborate with community colleges and public and private universities and colleges to provide enhanced educational opportunities for students.

NC Public Schools Will Be Governed andSupported By 21st Century Systems.

• Processes are in place for financial planning and budgeting that focus on resource attainment and alignment with priorities to maximize student achievement.

• Twenty-first century technology and learning tools are available and are supported by school facilities that have the capacity for 21st Century learning.

• Information and fiscal accountability systems are capable of collecting relevant data and reporting strategic and operational results.

• Procedures are in place to support and sanction schools that are not meeting state standards for student achievement.

Appendix D

Sample Race to the Top Omnibus Teacher and Principal Survey: Dimensions and Items

RttT Omnibus Survey Dimensions and Items Samples (Henry et al., 2014, p. 36-43)

Dimension:
Commitment, Climate, & Culture

Items:
Principal Instructional Leadership
•Makes clear to the staff his/her expectations for meeting instructional goals
•Communicates a clear vision for our school
•Sets high standards for teaching
•Understands how student learn
•Sets high standards for student learning
•Presses teachers to implement what they have learned in PD
•Carefully tracks student overall academic progress
•Actively monitors quality of teaching

Respect for Leadership
•I really respect my principal as an educator
•I have confidence in my principal's ability to do the job
•OK in this school to discuss feelings, worries, & frustrations with the principal
•I trust the principal at his or her word

Evaluation of Teachers
•Evaluation process encourages teachers to reflect on practice
•Teachers use feedback from teacher evaluation system to improvement their teaching
•Teacher evaluation is fair
•Criteria on which I am evaluated are clear
•Teacher evaluation process encourages professional growth
•This year overall, I am satisfied with the teacher evaluation process

Distributed Leadership
•At my school there is/are . . . specific strategies to develop shared leadership between administrators and teachers
•. . . opportunities for educators
to take on new leadership roles
•. . . structure, time, and resources for administrators & teachers to participate in joint decision-making
•. . . principal gives a leadership role to teaches who have special knowledge & skills in key areas
•. . . people who
have special knowledge or skills are allowed to take the

initiative
- . . . principal encourages teachers to take on informal leadership roles
- . . . principal has defined clear roles & responsibilities for dept chairs/team leaders

- . . . department chairs/team leaders are effective in their roles
- . . . principal likes it when teachers take initiative to deal with problems we face

Team Orientation
- At my school . . . the faculty has an effective process for making group decisions
- . . . teachers take an active role in school-wide decision making
- . . . useful information for solving problems is readily available to teachers
- . . . we work together as a team
- . . . we turn to each other for consultation & advice
- . . . we meet together to address challenges and solve problems
- . . . the faculty has an effective process for solving problems

Establishing Goals, Standards and Accountability; and Forging Bonds & Mobilizing Engagement
- Faculty & leadership have shared vision
- School improvement team provides effective leadership
- Share a high level of investment in what we are here to do
- Feel a strong sense of meaning & purpose in our work
- Share a common belief in the importance of our work

Safe, Orderly, & Caring Environment
- How safe do you feel. . . in the school building during school hours
- . . . in the school building after school hours
- . . . on school grounds/campus
- . . . in the neighborhood of the school
- To what extent have following been a problem this school year . . . physical conflict among students
- . . . vandalism of school property
- . . . student alcohol use
- . . . student tobacco use
- . . . student possession of weapons
- . . . verbal abuse of teachers
- . . . physical abuse of teachers
- . . . racial tensions

Teacher-Student Relationships
- I share an affectionate, warm relationship with most of the students in my classroom

•Most of the students in my classroom . . .will seek comfort from me if they are upset

•. . . value their relationships
with me

•. . . spontaneously share information about themselves with me

•. . . have a relationship with me that extends beyond academics

•. . . Openly share their feelings and experiences with me

Classroom Management

•I am effective at . . . calming a student who is disruptive or noisy

•. . . controlling disruptive behavior in the classroom

•. . . getting students to follow classroom rules

•. . . establishing a classroom management system with my students

•In my classroom . . . I structure instruction to encourage good behavior

•. . . desks and tables are arranged in a manner that encourages good
behavior

•. . . I have clearly posted rules of good behavior

•. . . there are
clear consequences for misbehavior

•. . . I am consistent at expecting good behavior, enforcing rules, and
disciplining bad behavior

•. . . I have clearly posted rules of good behavior

•. . . there are clear consequences for misbehavior

•. . . I am consistent at expecting good behavior, enforcing rules, and
disciplining bad behavior

Academic Rigor

•At my school . . . I stress to students the importance of trying hard and
persisting through academic challenges

•. . . I expect students to give their best effort

•. . . I set high standards for students' classroom performance

•. . . I have high expectations of my students

•. . . I assign work that is intellectually stimulating

•. . . I let students know when they are doing less than their best work

•. . . I encourage students to think they can do better in their
school work

•. . . I ask questions that are intellectually stimulating

•. . . I assign work that challenges my students

•. . . I let students know how they can improve their classroom performance

•. . . I let students know when they are doing their best work

Dimensions: Knowledge & Skills

Items:
Quality of PD

•At my school . . . teachers are left completely on their own to seek out PD (reverse scored, agree strongly=very negative, etc.)

•. . . most PD Topics offered once & not followed up (also reverse scored)

•Overall, my PD experiences this year . . . have been sustained rather than short term

•. . . coherently focused rather than unrelated

•. . . included enough time to think carefully about, try, and evaluate new ideas

•PD at my school has prepared teachers to . . . implement the Standard Course of Study

•. . . integrate computers & technology into lessons

•. . . develop warm relationships with students

•. . . implement good behavior management

•. . . use data to tailor instruction to students' needs

Alignment of PD

•At my school, most of what I learned in PD addresses the needs of the students in my classroom

•. . . Overall, my PD experiences this year have . . . been closely connected to my school's improvement plan school

•. . . included opportunities to work productively with
colleagues in my
school

•. . . included opportunities to work productively with teachers from
other schools

Dimension: Structures & Support for Instruction

Items:
Principal Effectiveness

•Principal at my school . . . has helped strengthen teachers' understanding of the Standard Course of Study (SCOS)

•. . . actively encourages teachers to use a pacing guide

•. . . makes sure
curriculum is clearly defined from grade to grade

•. . . actively encourages teachers to review curricular materials in light of SCOS

•. . . meets individually with teachers to discuss student progress
Common Core

•I fully understand the content knowledge & skills needed to master the subject(s) I teach

•I have the knowledge & skills I need to help my students accomplish what is expected of them in the state standards

•I feel well prepared to help students accomplish what is expected of them in the state standards

•Leaders in my school . . . discuss the state standards with teachers

•. . . provide opportunities for teachers to discuss understandings of the state standards across grades and content areas

•Teachers in my school . . . coordinate curriculum across grade levels according to the state standards

•. . . are clear about learning goals based on the state standards

•. . . have a clear understanding of how the state standards fit together across grade levels

•. . . focus more on teaching the state standards than "teaching to the test"

•. . . work together on developing/revising instructional units around state standards

•. . . are prepared to help students meet the state standards

Program Coherence

•At my school . . . once we start a new program, we follow up to make sure that it's working

•. . . we have so many different programs . . . that I can't keep track of them all (reverse scored: agree strongly=very negative)

•many special programs come and go

•. . . you can see real continuity from one program to another.

•. . . curriculum, instruction, and learning materials are well coordinated across the different grade levels.

•. . . there is consistency in curriculum, instruction, and learning materials in the same grade level.

Principal Instructional Leadership

•Principal at this school . . . makes clear to the staff his or her expectations for meeting instructional goals

•. . . communicates a clear vision for our school

•. . . sets high standards for teaching

•. . . presses teachers to implement what they have learned in professional development

•. . . carefully tracks student overall academic progress

•. . . actively monitors the quality of teaching in this school

Evaluation of Teachers

•At my school . . . the evaluation process encourages teachers to reflect on their instructional practice

•. . . teachers use feedback from the teacher evaluation system to improve their teaching

•. . . teacher evaluation is fair

•. . . the criteria on which I am evaluated are clear

•. . . the teacher evaluation process encourages professional growth

Teacher-Teacher Trust
- At my school . . . teachers feel respected by other teachers
- . . . teachers trust each other
- . . . it's okay for teachers to discuss feelings, worries, & frustrations with other teachers
- . . . teachers respect other teachers who take the lead in school improvement efforts
- . . . teachers respect those colleagues who are expert at their craft

Teacher Knowledge Sharing
- How frequently do teachers in your school . . . share ideas on teaching
- . . . share and discuss student work
- . . . discuss particular lessons that were not very successful
- . . . share and discuss research on effective teaching methods
- . . . share and discuss research on effective instructional practices for English language learners
- . . . explore new teaching approaches for under-performing students

Collective Responsibility
- How many teachers in your school . . . share a vision of good teaching
- . . . engage in systematic analysis of their teaching practices
- . . . feel responsible to help each other do their best
- . . . feel responsible for helping students develop self-control
- . . . feel responsible when students in this school fail

Data-driven Instruction
- At my school . . . teachers collect and use data to improve their teaching
- . . . we have made changes designed to better meet the needs of its diverse student body
- . . . teachers are engaged in systematic analysis of student performance data
- . . . assessment of student performance leads to changes in our school's curriculum that are consistent with state standards
- . . . useful information to make informed decisions is readily available to teachers (e.g., about student performance, resources, community satisfaction)
- . . . we use assessment data to evaluate teachers' instructional practices
- . . . the whole school examines gaps in the achievement of students by grade level
- . . . teachers meet regularly to review student performance in order to adjust their practices
- . . . we use a variety of assessment strategies to measure student progress

Level of Preparation for Tests
- This school year . . . I covered the material required by the State Standard

Course of Study
- . . . prepared my students for their EOG/EOC exams
- . . . prepared my students to move to the next level of schooling

Teacher Self-Efficacy
- If I try really hard, I can get through to even the most difficult students
- I am good at helping all the students in my classes make significant improvement
- I am certain that I am making a difference in the lives of my students
- I can deal with almost any learning problem

Teaching Practice
- Students in my classroom frequently . . . review and discuss the work of other students
- . . . explain their reasoning to the class
- . . . discuss ideas for a sustained period
- . . . work on a group project that extends for several days
- . . . reflect on their work and set future learning goals
- . . . ask probing questions about subject matter
- . . . reflect apply what they have learned to new questions, situations, and subjects
- . . . reflect on their own progress
- . . .express their own ideas about subject matter
- . . . provide constructive feedback to other students

Appendix E

Consent for Use of and Discussion about TEP (modified) with Dr. Duke

Email Correspondence Requesting Dr. Duke's Consent for Use of and Discussion about
the TEP (modified)

Teacher Evaluation Profile

Kim Case <kcase@█████████.com>   Tue, Mar 15, 2016 at 3:20 PM
To: dld7g@virginia.edu

Dr. Duke,
Currently, I am a doctoral candidate at Gardner-Webb University in North Carolina and am interested
in replicating a study completed by yourself and Richard Stiggins, which used the Teacher Evaluation
Profile questionnaire. I plan to examine educator's perceptions of the latest revision (July 2015) of the
North Carolina Teacher Evaluation Process (NCTEP) and will need to make slight modifications to the
survey to reflect the North Carolina's evaluation system.

May I have your permission to use the TEP with modifications, and if so, is the instrument available in
an electronic format?

I do have the original version of the TEP from *The Case for Commitment to Teacher Growth:
Research on Teacher Evaluation, 1988.* However, the document states that a revised version is
available. I investigated with Education Northwest and was referred to you.

Your assistance with my research is greatly appreciated.

Sincerely,
Kim Case

Duke, Daniel L. (dld7g) <dld7g@eservices.virginia.edu>   Tue, Mar 15, 2016 at 5:40
                                                                              PM
To: Kim Case <kcase@█████████>

> Kim: You have my permission to use the version of the TEP in THE CASE FOR
> COMMITMENT TO TEACHER GROWTH. Best of luck, Dan Duke, Professor of Educational
> Leadership, University of Virginia

Kim Case <kcase@█████████>   Wed, Mar 16, 2016 at 9:10
                                                                              AM
To: "Duke, Daniel L. (dld7g)" <dld7g@eservices.virginia.edu>

> Dr. Duke,
> Thank you for your prompt reply. In using the original version, I will need to eliminate or modify
> questions #53 and #54 since these are not applicable to N.C. Also, I plan to conduct the survey
> online and will need to adjust the directions to accommodate for this. May I have your permission
> to make slight revisions to the questionnaire to make adjustments for North Carolina's system and
> for online administration?
>
> I will gladly submit my revisions for your approval prior to conducting the survey. Kim

Duke, Daniel L. (dld7g) <dld7g@eservices.virginia.edu>   Wed, Mar 16, 2016 at 9:34 AM

To: Kim Case <kcase@

Kim: You have my permission to modify the TEP as noted in your email of March 16, 2016. I look forward to receiving your adjusted TEP. Daniel L. Duke

From: Kim Case Mon, Apr 11, 2016 at 11:35 AM
To: "Duke, Daniel L. (dld7g)" <dld7g@eservices.virginia.edu>
CC: Dr. Stephen C. Lawsslaws@gardner-webb.edu

Dr. Duke,

Thank you for your prior permission to use the Teacher Evaluation Profile (TEP) Questio
nnaire
as published in the work presented by yourself and Dr. Richard Stiggins in *The Case for
Commitment toTeacher Growth:*
Research on Teacher Evaluation, 1988, as a component of my
dissertation research.

As requested, I have attached a copy of my TEP (modified) for your review and
approval. To align the instrument with the North Carolina Teacher Evaluation
Process (NCTEP), 2015, which reflects recent legislative and curriculum changes,
it was necessary for me to modify the TEP more than I originally thought would be
necessary. Most changes were needed in Part C: Attributes of
Evaluation Procedures and Part E: Attributes of
the Evaluation Context. Should you have questions or concerns
about my revisions, I will be glad to revisit my work and re-submit for your review.

I have included my Dissertation Chair, Dr. Stephen Laws, from Gardner-Webb
University,
in this email as his help is critical as I move forward with my doctoral work. I have left
comments on the document that are directed to Dr. Laws.
You are welcome to respond as well.

Your assistance is much appreciated.

Sincerely,

 Kim Case

Duke, Daniel L. (dld7g) <dld7g@eservices.virginia.edu>   Tue, Apr 12, 2016 at 11:12 AM
To: Kim Case <kcase@ ▮▮▮▮▮▮▮▮▮▮ >

Dear Kim:

 Thank you for sharing your modified TEP. You've done a fine job aligning it with the
NCTEP. Of course, our validation process for the original TEP would not apply to your
changes, but that shouldn't be a major issue. My only concern involves the three impact
questions (4, 5, and 6). The items do not discriminate between a positive and a negative
impact. Either could occur. I understand that you get a general response in question 3, but
there still is the possibility of a negative impact on attitudes, for example, and a positive
impact on understanding of the teaching process. I wish you well on your doctoral work and
look forward to learning what you find. All the best, Daniel L. Duke

Teacher Evaluation Profile (TEP) Questionnaire - modified

Kim Case <kcase@ ▮▮▮▮▮▮▮▮▮ >   Tue, Apr 12, 2016 at 1:48 PM
To: "Duke, Daniel L. (dld7g)" <dld7g@eservices.virginia.edu>

Dr. Duke,
Thank you for your prompt reply. I will be proposing on Tuesday, April 19, and appreciate having
your information to share with the committee.

As you may have noted in the TEP (modified), I am planning to do follow-up interviews. Perhaps I
can gain a better understanding of the positive or negative impact from that data.

Are there any changes to the impact questions that you would suggest, or are you aware of other items
that have been used in other studies to determine whether the impact of the evaluation process is
positive or negative?

I believe my literature review has been fairly extensive, but I have found relatively little on teachers'
perceptions of the evaluation process since the early 2000s. There are studies about the inclusion of
VAMs that are more recent, but since I am interested in the holistic evaluation process, these studies
don't align well with mine.

Any advice is appreciated.

Sincerely,
Kim

Teacher Evaluation Profile (TEP) Questionnaire - modified

Duke, Daniel L. (dld7g) <dld7g@eservices.virginia.edu>

Tue, Apr 12, 2016 at 3:48 PM

To: Kim Case <kcase@███████████>

Kim:

You're certainly right about the lack of recent research on teachers' perceptions of teacher evaluation. Have you checked websites for the NEA and the AFT? In light of your post survey interviews, I think you'll be able to determine whether specific impacts have been positive or negative. It would be interesting if individuals differentiated their judgments of impact, rather than being all negative or all positive. It was interesting that you elected not to come right out and ask respondents if the NCTEP, in their judgment, had a positive impact on student achievement. Good look with your proposal defense, Dan Duke

Appendix F

Teacher Evaluation Profile (TEP) - modified

Teacher Evaluation Profile (TEP) Questionnaire (revised to reflect the North Carolina Teacher Evaluation Process (NCTEP)

This questionnaire has been designed to allow you to describe your experience with the **North Carolina Teacher Evaluation Process (NCTEP).** Your responses will be combined with those of other teachers to yield a clearer picture of teachers' perceptions of the NCTEP's impact on their own professional growth and teacher effectiveness, as well as how the evaluation process impacts their attitudes toward teaching. This research will assist in determining if and how the NCTEP is accomplishing its intended purpose: "to assess the teacher's performance in relation to the North Carolina Professional Teaching Standards and to design a plan for professional growth." (North Carolina Teacher Evaluation Process manual, revised July 2015, pg. 4)

This instrument has been designed to be comprehensive in scope and requires thoughtful responses. Please read and follow directions carefully. Your responses to this survey will remain anonymous. Responses will not be analyzed by the school where you are employed or connected to your school administrator in any way.

The form will take an estimated twenty to thirty minutes to complete. Your prompt response is needed. Response data will be collected in May 2016.

At the end of the survey, you will be asked if you are willing to participate in a follow-up interview. If so, you will provide a contact email or phone number. Your answers to this survey and your contact information will be kept confidential by the researcher. Your willingness to participate in this second phase of research is appreciated.

Definition of the Teacher Evaluation Process
(as outlined by the North Carolina Teacher Evaluation Process (NCTEP) manual, revised July 2015)

North Carolina GS 115c-333.1(a) requires that all teachers licensed by NCDPI adhere to specific evaluation guidelines. These requirements consist of:
- An annual evaluation process, based on one of three evaluation cycles (Comprehensive, Standard, or Abbreviated);
- The completion of the eight components of the NCTEP process. (The NCTEP components include training, orientation, self-assessment, observations with pre- and post-conferencing, a summary evaluation conference conducted by the principal and summative teacher rating, and professional development plans.);
- Inclusion of Standard 6 in the overall teacher evaluation. Ratings for Standard VI are not determined by the evaluator. Standard 6 is determined by a student growth value as calculated by the statewide growth model for educator effectiveness.

As reported on NCDPI's Educator Effectiveness Model webpage, an electronic data collection system, the North Carolina Educator Evaluation System (NCEES) is used during the North Carolina Teacher Evaluation Process. NCEES houses all evaluation information on a web-based platform. (Educator Effectiveness Model, http://www.ncpublicschools.org/effectiveness-model/)

Specific Instructions:

As you answer the following questions, consider that the North Carolina Evaluation Process (NCTEP) encompasses all elements listed in the prior definition of the teacher evaluation process.

Consider your most recent, completed annual evaluation cycle in the North Carolina Teacher Evaluation Process (NCTEP) when answering the survey questions.

Consent Form

The purpose of this research project is to determine teachers' perceptions of the impact that the North Carolina Teacher Evaluation Process (NCTEP) has on teacher effectiveness and professional growth. This is a research project being conducted by Kim Case, a doctoral student, at Gardner-Webb University. You are invited to participate in this research project because you are a North Carolina teacher.

Your participation in this research study is voluntary. You may choose not to participate. If you decide to participate in this research, you may withdraw at any time. If you decide not to participate in this study or if you withdraw from participating at any time, you will not be penalized.

The procedure involves completing an online survey that will take approximately 30 minutes. Your responses will be confidential. If you choose to participate in a follow-up interview session with the researcher by providing contact information, your contact information will be kept confidential and coded to protect your anonymity. The survey questions will be about your experiences with the North Carolina Teacher Evaluation Process (NCTEP).

The researcher will be diligent in keeping your information confidential. All data will be stored in a password protected electronic format. The results of this study will be used for scholarly purposes only.

If you have any questions about the research study, please contact Kim Case at kcase@████████.com. This research has been reviewed according to Gardner-Webb University IRB procedures for research involving human subjects.

* 1. Electronic Consent: Please select your choice below.
Clicking on the "agree" button below indicates that:
- **you have read the above information;**
- **you voluntarily agree to participate;**
- **you are at least 18 years of age.**

If you do not wish to participate in the research study, please decline participation by clicking on the "disagree" button.

○ Agree

○ Disagree

Qualifying Questions

2) I am a teacher, licensed by the Public Schools of North Carolina, and am subject to evaluation through the NCTEP.
   a. Yes
   b. No
      (*If you responded "No" to Question #1, you have completed this survey. Thank you for your willingness to participate.*)

3) My most recent, completed annual evaluation cycle, which concluded in a summary evaluation conference, was for the:
   a. 2015-16 school year;
   b. 2014-15 school year;
   c. Other *(Give most recent evaluation year.)*
   d. I have not completed an annual evaluation cycle.
      (*If you responded "No" to Question #1, you have completed this survey. Thank you for your willingness to participate.*)

Overall Experience with the NCTEP

4) Rate the *overall quality* of the entire evaluation process.
   (0=poor quality; 9=high quality)
   Poor Quality   0   1   2   3   4   5   6   7   8   9   High Quality

5) Rate the *impact* of the evaluation process *on your attitudes about teaching.*
   (A strong impact rating (9) would reflect a profound change in how you feel about the content you teach, your students, and/or yourself as a teacher.)
   No Impact   0   1   2   3   4   5   6   7   8   9   Strong Impact

6) Rate the *impact* of the evaluation process *on your teaching behaviors and strategies.*
   (A strong impact (9) would reflect major changes in your instructional behavior, classroom management strategies, evaluation practices, and/or other observable dimensions of your teaching.)

   No Impact   0   1   2   3   4   5   6   7   8   9   Strong Impact

7) Rate the *impact* of the evaluation process *on your understanding of the teaching/learning process.*
   (A strong impact (9) would reflect a change in your ability to account for your effectiveness (or lack thereof), explain the reasons for your instructional decisions, and/or better understanding of student needs or behavior.)

Part A: Teacher Attributes

Describe these attributes of you as a teacher.

8) Rate the strength of your professional expectations of yourself.
   I demand little. A   B   C   D   E   I demand a great deal.

9) In my last completed evaluation cycle, I received an overall status rating of:
    a. In Need of Improvement
    b. Effective
    c. Highly Effective
    d. Did not receive a status rating

Describe your interpersonal manner:

10) Orientation to risk taking
    I avoid risks. A   B   C   D   E   I take risks.

11) Orientation to change
    I'm relatively slow to change. A   B   C   D   E   I'm relatively flexible.

12) Orientation to experimentation in classroom
    I don't experiment. A   B   C   D   E   I experiment frequently.

13) Openness to criticism
    I'm relatively closed. A   B   C   D   E   I'm relatively open.

14) Knowledge of technical aspects of teaching
    I know a little. A   B   C   D   E   I know a great deal.

15) Knowledge of subject matter
    I know a little. A   B   C   D   E   I know a great deal.

Describe your teaching and evaluation experience:

16) Total years in teaching:
    a. 0 to 1 year
    b. 2 to 5 years
    c. 6 to 10 years
    d. 11 to 19 years
    e. 20 or more years

17) Grade level taught at the time of most recent, completed annual evaluation cycle:
    a. Grades K to 2
    b. Grades 3 to 5
    c. Grades 6 to 8
    d. Grades 9 to above
    e. Other or multiple grade levels
       (List grade level(s) taught: _____)

18) Experience with the **NC Standard Course of Study**, *prior to implementation of the Common Core Standards*
    a. No experience
    b. 1 to 2 years
    c. 3 to 5 years
    d. 6 to 10 years
    e. 11 or more years

19) Experience with the **NC Standard Course of Study**, _since the implementation of the Common Core Standards_
    a. No experience
    b. 0 to 1 year
    c. 2 to 3 years
    d. 4 to 5 years
    e. 6 or more years

20) I have experience with the _North Carolina Teacher Performance Appraisal Instrument, revised (TPAI-R)_, which was used for the NC teacher evaluation process prior to 2008.
    a. Yes
    b. Yes, but I don't recall details of the process.
    c. No

21) If you replied "Yes," rate your experience with the _TPAI-R_.
    Waste of time   A   B   C   D   E   Helpful

Part B: Evaluator Attributes

Describe your perceptions of the person who conducted your most recently completed, annual summary evaluation conference.

22) Credibility as a source of feedback
    Not credible   A   B   C   D   E   Very credible

23) Working relationship with you
    Adversary   A   B   C   D   E   Helper

24) Level of trust
    Not trustworthy   A   B   C   D   E   Trustworthy

25) Interpersonal manner
    Threatening   A   B   C   D   E   Not threatening

26) Temperament
    Impatient   A   B   C   D   E   Patient

27) Flexibility
    Rigid   A   B   C   D   E   Flexible

28) Knowledge of technical aspects of teaching
    Not knowledgeable   A   B   C   D   E   Knowledgeable

29) Capacity to demonstrate or model needed improvements
    Low   A   B   C   D   E   High

30) Familiarity with your particular classroom
    Unfamiliar   A   B   C   D   E   Very familiar

31) Experience in the classroom in general
Little   A   B   C   D   E   A great deal

32) Usefulness of suggestions for improvement
Useless   A   B   C   D   E   Useful

33) Persuasiveness of rationale for suggestions
Not persuasive   A   B   C   D   E   Very persuasive

Part C: Attributes of Evaluation Procedures

Describe these attributes of the information gathered on your performance during your most recent evaluation.

How were the dimensions of your teaching (standards) to be evaluated/addressed?

34) Were standards communicated to you?
Not at all   A   B   C   D   E   In great detail

35) Were standards clear to you?
Vague   A   B   C   D   E   Clear

36) How accurately was your teacher effectiveness measured as based on your annual summative teacher rating? (Standards 1-5)
Not accurate   A   B   C   D   E   Very Accurate

37) How accurately was your teacher effectiveness measured as based on student growth measures in the Education Value-Added Assessment System (EVAAS)? (Standard 6)

Not accurate   A   B   C   D   E   Very Accurate

To what extent were the following sources of performance information tapped as part of the evaluation?

38) Observation of your classroom performance
Not considered   A   B   C   D   E   Used extensively

39) Examination of classroom artifacts
Not considered   A   B   C   D   E   Used extensively

40) Examination of student achievement
Not considered   A   B   C   D   E   Used extensively

41) Progress toward Professional Development Plan goals
Not considered   A   B   C   D   E   Used extensively

42) In my most recent, completed annual evaluation cycle, my observations included:
   a. Three full class periods;
   b. One full class period and two observations that were less than an entire class period;
   c. Two observations that were less than an entire class period;
   d. No observations
   e. Other (Please describe.)

43) During my last completed evaluation cycle, I was observed by: *(Choose all that apply.)*
   a. Principal only
   b. Principal and Assistant Principal(s)
   c. Other observers. (Please give titles.)
   d. Not observed

Part D: Attributes of Feedback Received

Describe these attributes of the feedback you received.

44) Amount of information received
   None   A   B   C   D   E   Great deal

45) Frequency of formal feedback
   Infrequent   A   B   C   D   E   Frequent

46) Frequency of informal feedback
   Infrequent   A   B   C   D   E   Frequent

47) Depth of information provided
   Shallow   A   B   C   D   E   In-depth

48) Quality of the ideas and suggestions contained in the feedback
   Low   A   B   C   D   E   High

49) Specificity of information provided
   General   A   B   C   D   E   Specific

50) Nature of information provided
   Judgmental   A   B   C   D   E   Descriptive

51) Timing of feedback
   Delayed   A   B   C   D   E   Immediate

52) Feedback focused on teaching standards
   Ignored them   A   B   C   D   E   Reflected them

Part E: Attributes of the Evaluation Context

Describe these attributes of the evaluation context.

53) Use of the NCEES online platform for data collection enhanced the evaluation process.
None   A   B   C   D   E   Great deal

What resources are available for professional development?

54) What level of the NC educational system provided the highest amount of support for professional development to assist in reaching Professional Development Plan goals?

    a. State (NCDPI)
    b. District
    c. School

55) Amount of professional development support provided by this entity.
None   A   B   C   D   E   Great deal

56) Time allotted during the school day for professional development related to Professional Development Plan goals.
None   A   B   C   D   E   Great deal

How were federal/state/district values and policies expressed in evaluation?

57) Clarity of policy statements regarding purpose for evaluation
Vague   A   B   C   D   E   Clear

58) Intended role of evaluation
Teacher accountability   A   B   C   D   E   Teacher Growth

59) Impact of federal initiatives, such as Race to the Top, on the evaluation process
None   A   B   C   D   E   Great deal

60) Impact of state law and policies on evaluation process
None   A   B   C   D   E   Great deal

Part F: Other Items for Discussion

61) Are there other dimensions of you as a teacher, the nature of the performance data collected, the nature of the feedback, the evaluation context, or other factors that you think are related to the success(or lack of success) of your past teacher evaluation experiences that should be included in the above list? If so, please specify.

62) If you are willing to participate in a follow-up interview, please provide a contact e-mail or phone number. Your contact information will be kept confidential by the researcher.

    a. Yes, I am willing to participate.
    b. I can be contacted at: _____.

You have completed the Teacher Evaluation Profile Questionnaire. Your participation is much appreciated.

Appendix G

Consent for Use of CERE-NC's Interview Protocol Questions from C. Davis

CERE-NC Research

Kim Case <kcase@█████████████>Fri, Apr 8, 2016 at 9:39 AM
To: cnrichar@email.unc.edu

Ms. Davis,
I am currently completing doctoral work through Gardner-Webb University. My topic centers
around teachers' perceptions of the impact of the NCTEP on teacher effectiveness and professional growth.
As part of my research, I plan to utilize a mixed methods approach and
would like to include some of the survey questions from the Teacher Interview Protocol described in the Teacher and Principal Perceptions of the North Carolina Educator Evaluation System: Final Evaluation Report on pgs. 32 - 35 in my work.

May I have permission to use these questions?   Thank you for your assistance.

Kim Case

Davis, Cassandra Richards <cnrichar@email.unc.edu>   Fri, Apr 8, 2016 at 10:16 AM
To: Kim Case <kcase@█████████████>

   Hello Ms. Case,

   You have my permission to use our survey questions. Good luck with your
   dissertation and let me know if you have any questions.

   Best,
   Cassandra

   --
   Cassandra R. Davis, Ph.D.
   Research Associate
   Education Policy Initiative at Carolina (EPIC)
   UNC Public Policy
   314 Cloister Court
   Chapel Hill, NC 27514
   (o): 919.962.1338
   cnrichar@email.unc.edu

Appendix H

Researcher-developed Interview Items for Phase Two: Follow-up Interviews

Researcher-developed Interview Items for Phase Two: Follow-up Interviews

Thank you for participating in this interview. During the interview, I will be recording our conversation for transcription. You may request a copy of the transcription notes.

- May I have your permission to record our conversation?
- Would you like a copy of the notes?
  (So that your responses will remain anonymous during the interview, I will get your contact information after I have stopped the recording.)

Turn on recording.

As you may recall from the online survey, this study hopes to provide insights into how the North Carolina Teacher Evaluation Process impacts teachers' perceptions of professional growth, teacher effectiveness, and attitudes toward teaching.

1) To begin, please tell me your years of teaching experience.

2) What grade level did you teach during your last evaluation cycle?

3) Please give a description of your experiences with the North Carolina Teacher Evaluation Process.

4) What aspects of the evaluation process have been most helpful in improving your professional practice?

- In what ways have these elements helped you grow as a teacher?
5) What aspects have been the least helpful in your professional growth?

- How could these elements be improved?
6) On the survey, the characteristics of the evaluation process are categorized as five attributes. These are: Teacher, Evaluator, Evaluation Process, Feedback and Context of the Evaluation.

- Which of these categories is the most significant to your professional growth?
- Please explain why.
7) Next, consider the relationship between evaluation and teacher effectiveness.

- How has the evaluation process impacted your effectiveness as a teacher?
- Are there elements or attributes of the process that contribute more to your effectiveness than the others?
- What do you think makes the difference in their impact?

8) Does the evaluation process impact your attitude toward teaching?

- Why **or** why not?
- Describe how your attitude is impacted.

Next, I'd like to get your ideas about some of the survey results.

9) Survey data showed that there was not a relationship between teacher attributes, such as flexibility and expectations of self, and teacher ratings of the quality of the evaluation process or to its impact on attitudes toward teaching, professional growth or teacher effectiveness.

- What do you think that this indicates?

10) On the hand, the context of the evaluation, which includes purposes of the evaluation and professional development connected to PDP goals, had a significant relationship to the quality of the evaluation and the impact on attitudes, professional growth and teacher effectiveness.

- What do you think this data might be describing?

11) In the survey, teachers were asked to qualify whether the intended role of the evaluation was primarily teacher accountability or teacher growth.

Results from the survey showed that slightly more than a third of the respondents felt that there was a balance between accountability and growth. Less than a fourth (about 23%) of the respondents felt that the role was more toward accountability, while the majority of responses (about 41%) leaned more toward teacher growth.

- What insights can you provide about this distribution of responses?

12) In conclusion, is there anything that you would like to add that I may not have asked you?

Thank you for participating in this interview. I appreciate your assistance.

Contact Information:

_____

_____

_____

_____

Appendix I

Frequency Tables for TEP (modified)
Q4 – Q60

Frequency Tables for TEP (modified): Q4 – Q60

Q4 - Overall Quality of Evaluation Process

	Frequency	Valid Percent
Poor Quality	8	4.1
1	4	2.0
2	8	4.1
3	16	8.1
4	28	14.2
5	40	20.3
6	32	16.2
7	34	17.3
8	14	7.1
High Quality	13	6.6
Total	197	100.0

Q5- Impact on Attitudes about Teaching

	Frequency	Valid Percent
No Impact	16	8.1
1	10	5.1
2	13	6.6
3	24	12.2
4	27	13.7
5	33	16.8
6	27	13.7
7	20	10.2
8	13	6.6
Strong Impact	14	7.1
Total	197	100.0

Q6) Impact on Teaching Behaviors and Strategies

	Frequency	Valid Percent
No Impact	16	8.2
1	11	5.7
2	10	5.2
3	22	11.3
4	28	14.4
5	44	22.7
6	27	13.9
7	22	11.3
8	8	4.1
Strong Impact	6	3.1
Total	194	100.0

Q7 Impact on Understanding of the Teaching/Learning Process

	Frequency	Valid Percent
No Impact	27	13.8
1	10	5.1
2	21	10.7
3	14	7.1
4	29	14.8
5	33	16.8
6	26	13.3
7	23	11.7
8	6	3.1
Strong Impact	7	3.6
Total	196	100.0

Teacher Attributes

Q8 Expectations of Self

	Frequency	Valid Percent
3	7	3.6
4	54	28.1
I demand a great deal.	131	68.2
Total	192	100.0

Q9 Status Rating

	Frequency	Valid Percent
In Need of Improvement	2	1.0
Effective	64	33.3
Highly Effective	113	58.9
No Status Rating	10	5.2
No Response	3	1.6
Total	192	100.0

Q10 Risk-taking

	Frequency	Valid Percent
Avoid risks.	3	1.6
2	17	8.9
3	71	37.0
4	53	27.6
Take risks.	48	25.0
Total	192	100.0

Q11 Orientation to Change

	Frequency	Valid Percent
Slow to change	3	1.6
2	7	3.6
3	38	19.8
4	44	22.9
Flexible	100	52.1
Total	192	100.0

Q12 Experiment in Class

	Frequency	Valid Percent
I don't experiment.	1	.5
2	6	3.1
3	37	19.3
4	76	39.6
Experiment frequently	72	37.5
Total	192	100.0

Q13 Openness to Criticism

	Frequency	Valid Percent
2	4	2.1
3	41	21.4
4	70	36.5
Relatively Open	77	40.1
Total	192	100.0

Q14 Technical Knowledge

	Frequency	Valid Percent
Know a little.	1	.5
2	8	4.2
3	31	16.1
4	83	43.2
Know a great deal	69	35.9
Total	192	100.0

Q15 Subject Knowledge

	Frequency	Valid Percent
3	11	5.7
4	58	30.2
Knows a great deal.	123	64.1
Total	192	100.0

Q16 Experience (years)

	Frequency	Valid Percent
0-1 yr.	5	2.6
2-5 yrs.	26	13.5
6-10 yrs.	30	15.6
11-19 yrs.	68	35.4
20 yrs.or more	63	32.8
Total	192	100.0

Q17 Experience (grade level)

	Frequency	Valid Percent
PreK	3	1.6
K-2	21	11.0
3-5	30	15.7
6-8	51	26.7
9 and above	56	29.3
Other	30	15.7
Total	191	100.0

Q18 Experience (NCSCOS, pre-CCS)

	Frequency	Valid Percent
No experience	16	8.3
1-2 yrs.	16	8.3
3-5 yrs.	41	21.4
6-10 yrs.	35	18.2
11 yrs. or more	84	43.8
Total	192	100.0

Q19 Experience (NCSCOS, since CCS)

	Frequency	Valid Percent
No experience	7	3.6
0-1 yrs.	11	5.7
2-3 yrs.	43	22.4
4-5 yrs.	67	34.9
6 yrs. or more	64	33.3
Total	192	100.0

Q20 NC TPAI-R (Experience)

	Frequency	Valid Percent
Yes	82	42.9
Don't remember	53	27.7
No	56	29.3
Total	191	100.0

Q21 NCTPAI-R (Rating)

	Frequency	Valid Percent
0	85	45.0
Waste of time	11	5.8
2	12	6.3
3	40	21.2
4	19	10.1
Helpful	20	10.6
9	2	1.1
Total	189	100.0

Evaluator Attributes

Q22 Credibility - July 1, 2016

	Frequency	Valid Percent
Not credible	8	4.3
2	14	7.5
3	29	15.5
4	38	20.3
Very Credible	98	52.4
Total	187	100.0

Q23 Relationship to Teacher

	Frequency	Valid Percent
Adversary	6	3.2
2	11	5.9
3	28	14.9
4	46	24.5
Helper	97	51.6
Total	188	100.0

Q24 Trustworthiness

	Frequency	Valid Percent
Not Trustworthy	6	3.3
2	14	7.8
3	21	11.7
4	46	25.6
Trustworthy	93	51.7
Total	180	100.0

Q25 Interpersonal Manner

	Frequency	Valid Percent
Threatening	4	2.1
2	7	3.7
3	33	17.6
4	39	20.7
Non-threatening	105	55.9
Total	188	100.0

Q26 Temperament

	Frequency	Valid Percent
Impatient	5	2.7
2	9	4.8
3	32	17.2
4	35	18.8
Patient	105	56.5
Total	186	100.0

Q27 Flexibility

	Frequency	Valid Percent
Rigid	9	4.8
2	12	6.4
3	29	15.4
4	41	21.8
Flexible	97	51.6
Total	188	100.0

Q28 Technical Knowledge

	Frequency	Valid Percent
Not Knowledgeable	3	1.6
2	15	8.0
3	28	14.9
4	45	23.9
Knowledgeable	97	51.6
Total	188	100.0

Q29 Capacity to Model

	Frequency	Valid Percent
Low	17	9.1
2	20	10.7
3	33	17.6
4	55	29.4
High	62	33.2
Total	187	100.0

Q30 Familiar with Classroom

	Frequency	Valid Percent
Unfamiliar	16	8.5
2	19	10.1
3	30	16.0
4	52	27.7
Very Familiar	71	37.8
Total	188	100.0

Q31 Classroom Experience

	Frequency	Valid Percent
Little	12	6.4
2	18	9.6
3	44	23.4
4	45	23.9
A Great Deal	69	36.7
Total	188	100.0

Q32 Useful Suggestions

	Frequency	Valid Percent
Useless	15	8.0
2	23	12.2
3	33	17.6
4	45	23.9
Useful	72	38.3
Total	188	100.0

Q33 Persuasive Rationale

	Frequency	Valid Percent
Not Persuasive	8	4.3
2	23	12.2
3	43	22.9
4	59	31.4
Persuasive	55	29.3
Total	188	100.0

Attributes of the Evaluation Procedures
Q34 NCPTS – Communicated

	Frequency	Valid Percent
Not at all	6	3.2
2	21	11.4
3	64	34.6
4	51	27.6
In Great Detail	43	23.2
Total	185	100.0

Q35 NCPTS – Clarity

	Frequency	Valid Percent
Not at All	5	2.7
2	14	7.6
3	42	22.7
4	63	34.1
Clear	61	33.0
Total	185	100.0

Q36 Accuracy of Summative Rating (Stan. 1 – 5)

	Frequency	Valid Percent
Not Accurate	3	1.6
2	17	9.2
3	36	19.5
4	79	42.7
Very Accurate	50	27.0
Total	185	100.0

Q37 Accuracy of EVAAS Rating (Stan. 6)

	Frequency	Valid Percent
Not Accurate	22	11.9
2	28	15.1
3	56	30.3
4	50	27.0
Very Accurate	29	15.7
Total	185	100.0

Q38 Use of Classroom Observations

	Frequency	Valid Percent
Not Considered	4	2.2
2	11	6.1
3	52	28.7
4	66	36.5
Used Extensively	48	26.5
Total	181	100.0

Q39 Use of Classroom Artifacts

	Frequency	Valid Percent
Not Considered	29	15.8
2	26	14.2
3	55	30.1
4	49	26.8
Used Extensively	24	13.1
Total	183	100.0

Q40 Use of Student Achievement

	Frequency	Valid Percent
Not Considered	12	6.6
2	16	8.8
3	58	31.9
4	53	29.1
Used Extensively	43	23.6
Total	182	100.0

Q41 Examine PDP Progress

	Frequency	Valid Percent
Not Considered	10	5.5
2	22	12.0
3	46	25.1
4	70	38.3
Used Extensively	35	19.1
Total	183	100.0

Q42 Observation Type

	Frequency	Valid Percent
3 Full Class Periods	49	26.8
1 Full and 2 Partial Class Periods	68	37.2
2 Partial Class Periods	56	30.6
No Observations	2	1.1
Other	8	4.4
Total	183	100.0

Q43 Observer(s)

	Frequency	Valid Percent
Principal Only	81	44.5
Principal and Asst. Principal	68	37.4
Not Observed	1	.5
Other	32	17.6
Total	182	100.0

Attributes of Feedback Received

Q44 Amount of Information

	Frequency	Valid Percent
None	6	3.3
2	30	16.5
3	53	29.1
4	54	29.7
A Great Deal	39	21.4
Total	182	100.0

Q45 Frequency – Formal

	Frequency	Valid Percent
Infrequent	23	12.6
2	28	15.4
3	53	29.1
4	45	24.7
Frequent	33	18.1
Total	182	100.0

Q46 Frequency – Informal

	Frequency	Valid Percent
Infrequent	21	11.5
2	24	13.2
3	46	25.3
4	49	26.9
Frequent	42	23.1
Total	182	100.0

Q47 Depth of Information

	Frequency	Valid Percent
Shallow	18	9.9
2	30	16.5
3	55	30.2
4	41	22.5
In-depth	38	20.9
Total	182	100.0

Q48 Quality of Ideas

	Frequency	Valid Percent
Low	20	11.0
2	29	15.9
3	52	28.6
4	37	20.3
High	44	24.2
Total	182	100.0

Q49 Specificity of Information

	Frequency	Valid Percent
General	22	12.1
2	20	11.0
3	59	32.4
4	36	19.8
Specific	45	24.7
Total	182	100.0

Q50 Nature of Information

	Frequency	Valid Percent
Judgemental	3	1.6
2	17	9.3
3	61	33.5
4	48	26.4
Descriptive	53	29.1
Total	182	100.0

Q51 Timing of Feedback

	Frequency	Valid Percent
Delayed	16	8.8
2	13	7.1
3	39	21.4
4	55	30.2
Immediate	59	32.4
Total	182	100.0

Q52 Focused on Standards

	Frequency	Valid Percent
Ignored Them	6	3.3
2	19	10.6
3	50	27.8
4	54	30.0
Reflected Them	51	28.3
Total	180	100.0

Attributes of the Context of the Evaluation

Q53 Use of NCEES

	Frequency	Valid Percent
Not at All	32	18.0
2	26	14.6
3	57	32.0
4	35	19.7
A Great Deal	28	15.7
Total	178	100.0

Q54a PD Providers (PDP Goals)

	Frequency	Valid Percent
State (NCDPI)	32	100.0

Q54b PD Providers (PDP Goals)

	Frequency	Valid Percent
District	84	100.0

Q54c PD Providers (PDP Goals)

	Frequency	Valid Percent
School	125	100.0

Q55 Amount of PD Support

	Frequency	Valid Percent
None	4	2.2
2	23	12.7
3	56	30.9
4	52	28.7
A Great Deal	46	25.4
Total	181	100.0

Q56 Time for PD

	Frequency	Valid Percent
None	36	19.9
2	43	23.8
3	53	29.3
4	33	18.2
A Great Deal	16	8.8
Total	181	100.0

Q57 Clarity of Evaluation Policy

	Frequency	Valid Percent
Vague	17	9.4
2	19	10.6
3	67	37.2
4	33	18.3
Clear	44	24.4
Total	180	100.0

Q58 Intended Role of Evaluation

	Frequency	Valid Percent
Teacher Accountability	27	14.9
2	16	8.8
3	62	34.3
4	42	23.2
Teacher Growth	34	18.8
Total	181	100.0

Q59 Federal Impact

	Frequency	Valid Percent
None	30	16.7
2	30	16.7
3	56	31.1
4	37	20.6
A Great Deal	27	15.0
Total	180	100.0

Q60 State Impact

	Frequency	Valid Percent
None	17	9.4
2	23	12.8
3	52	28.9
4	44	24.4
A Great Deal	44	24.4
Total	180	100.0