Teachers' and Administrators' Perceptions of the Kansas Educator Evaluation Protocol

Brooke L. Brutto
B.S., Baker University, 2007
M.S., Baker University, 2010

Submitted to the Graduate Department and Faculty of the School of Education of Baker University in partial fulfillment of the requirements for the degree of Doctor of Education in Educational Leadership

_____
Susan K. Rogers, Ph.D.
Major Advisor

_____
Harold B. Frye, Ed.D.

_____
Jill Smith, Ed.D.

Date Defended: April 28, 2016

Abstract

The setting for this research study was a school district located in Johnson County, Kansas serving 5,452 students. Within the district, there were 11 schools consisting of seven elementary schools, three middle schools, and one high school. The sample for this study included 438 teachers and 32 administrators who were employed within the district during the 2015-2016 school year.

The purpose of this study was to determine teachers' and administrators' understanding of the differences among the teacher performance ratings and of the calculation of the final summative teacher evaluation rating and to determine teachers' and administrators' perceptions of whether the district is using the KEEP model. Additionally, the purpose was to determine whether school level and whether a teacher had been evaluated affected teacher understanding and perceptions and whether school level affected administrator understanding and perceptions.

The findings of the research study were mixed. Teachers and administrators mostly or completely understood the differences among the performance ratings of highly effective, effective, developing, or ineffective. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective were not affected by school level. On average, teachers who had been evaluated better understood the differences in the performance ratings than teachers who had not been evaluated. The study results also indicated that there were no differences between teachers' and administrators' understanding among the performance ratings. Teachers mostly or completely understood the calculation of the final summative teacher evaluation rating and their understanding of the final summative teacher evaluation rating

was not affected by school level. However, teachers who had been evaluated with the KEEP evaluation tool understood the calculation of the final summative rating better than those who had not yet been evaluated using the KEEP evaluation tool. Administrators mostly understood the calculation of the final summative teacher evaluation rating. On average, administrators understood the calculation of the final summative teacher evaluation rating better than teachers understood it. Teachers agreed that the district was using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions and their perceptions were not affected by school level. However, teachers' perceptions that the district was using the KEEP Model for personnel decisions were affected by school level. Elementary teachers' perceptions were significantly lower than middle school teachers perceptions were in regards to how the KEEP Model was used for personnel decisions. Administrators agreed that the district was using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions. There were no differences between teachers' and administrators' perceptions that the district was using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions. It can be concluded that District G is in need of professional development for both teachers and administrators on the KEEP evaluation process. School districts, administrators, and teachers should analyze their understanding of the KEEP instrument.

To support highly effective educators in each classroom, it is imperative that the instrument used for evaluation by school district integrates instructional practice and feedback that is supported by research.

Dedication

This dissertation is dedicated to all the women who have supported, inspired, and challenged me in my life, career, and walk in faith. To my sister, who has kept me centered and faith driven. To my 98-year-old grandmother, from whom I inherited my stubbornness and drive. To my mother, who never lets me give up or walk away from something I truly want. To my best friend, who walked beside me as I questioned my perseverance and self-expectations.

Acknowledgements

With sincere gratitude, I would like to thank my mom for her support, words of encouragement, and assurance during the last three years. You have always believed in me and never questioned my ability, even when life threw challenges that were unexpected. Your words of love and belief allowed me to push through the obstacles and accomplish my biggest dreams.

Thank you to Peg Waterman for serving as my research analyst. Your patience and guidance through all my complicated questions have been a true blessing to me throughout this process. Thank you to Dr. Harold Frye for serving on my dissertation committee. I have always admired your authentic personality and ability to inspire future administrators. Also, a sincere appreciation is sent to Dr. Jill Smith for serving on my dissertation committee and being a mentor for me throughout my career. Because of your guidance when my career challenged me, and the positive perspective you offered in each situation, I hope to inspire others the way you have inspired me. Finally, I extend my most sincere appreciation to Dr. Susan Rogers. I would not have completed this goal without your patience, guidance, and advice. It was possible only because of you.

Table of Contents

List of Tables

Chapter One

Introduction

In 2012, the United States Department of Education began granting state waivers to the No Child Left Behind Act (National Education Association, 2010). The waiver required states to demonstrate a commitment to an assortment of school reforms (National Education Association, 2010). Among the school reforms needing action by states, was revamping teacher evaluations to include the academic performance of students (National Education Association, 2010). The core purposes of teacher evaluation or assessments should be to "strengthen the knowledge, skills, dispositions, and classroom practices of professional educators" (National Education Association, 2010, p. 2). In order to address these purposes, Kansas Statutes Annotated (KSA) 72-9003 has required all Kansas districts to implement "a written policy of personnel evaluation procedure in accordance with the law as outlined in KSA #72-9004, and file the same with the State Board" (Kansas State Department of Education (KSDE, 2014, p. 1). In 2013, KSDE created the Kansas Education Evaluation Protocol (KEEP) that school districts are encouraged to use. School districts that would like to submit their evaluation system for approval must meet the Kansas Educator Evaluation Guidelines established by the KSDE, as outlined in the Elementary and Secondary Education Act Flexibility Waiver Request (KSDE, 2014). During the 2014-2015 school year, all school districts in Kansas were required to implement an evaluation system that met the guidelines set by the KSDE (KSDE, 2014).

Throughout the 2014-2015 school year, KSDE held webinars and made presentations at various districts covering the topic of the Kansas Educator Evaluation

Protocol (KEEP) (KSDE, 2014). Webinars were specifically designed for teachers and administrators using the KEEP portal and uploading artifacts throughout the evaluation process (KSDE, 2014). In addition, KSDE created three different positions within the state department that are dedicated to the support of teachers and administrators who use the KEEP system within their district: Teacher Licensure and Accreditation Team Director, Assistant Director, and an Education Program Consultant (KSDE, 2014).

As the second full year of implementation ended in 2014-2015, the KSDE and Kansas school districts must continue to align professional development for teachers and administrators (KSDE, 2014). The main goal of professional development is to convey the objective of the six guidelines within the KEEP and support school community understanding of the calculation of the final summative rating (KSDE, 2014). KSDE created many resources that districts can access online to facilitate professional learning for KEEP.

Background

According to KSDE (2015), District G is located in Johnson County, Kansas and is home to 5,452 students. Within District G, there are 11 schools consisting of seven elementary schools, three middle schools, and one high school. During the 2015-2016 school year, 438 teachers were employed within the district and were evaluated using KEEP. The 2015-2016 school year was the second year that KEEP was used to evaluate teachers within the district.

KEEP was first piloted in 2011-2012 in 12 districts. In 2012-2013, 17 districts piloted KEEP within their school communities. Teacher unions and school boards reviewed the protocol and constructed feedback for the Kansas State Board of Education.

The first full year of implementation across Kansas was during the 2013-2014 school year. After the completion of the first full year of implementation, an annual report and study were constructed, and statewide user groups met quarterly with the state board council.

The KEEP Evaluation System and any other evaluation submitted to the state for approval met the following guidelines: "Used for Continual Improvement-Statement of Philosophy; Meaningfully Differentiates Performance; Based on Multiple Valid Measures; Evaluates Educators on a Regular Basis; Provides Useful Feedback; and, Used for Informed Personnel Decisions" (KSDE, 2014, pp. 2-4). Within the guidelines are also timelines for evaluations. During the first two consecutive years of employment, educators must be evaluated one time per semester. During the third and fourth years of employment, educators must be evaluated once by February 15 of each year. For the fifth year of employment and beyond, educators must be evaluated once every three years (KEEP, 2014).

District G for teachers implemented KEEP evaluations during the 2014-2015 school year. Administrators began training with district personnel in the summer of 2014. Throughout the school year, administrators attended KSDE webinars held at the district provided by the Education Program Consultant. Teachers and administrators attended professional development in August of 2014 and 2015 regarding the KEEP rubrics and portal. In October of 2015, teachers worked with their discipline colleagues to select Student Growth Measures. Administrators were also asked to provide goal-setting expectations and support for educators currently not on the evaluation cycle for the 2015-2016 school year.

Statement of the Problem

Danielson (2011) stated that she believed an evaluation system should include a consistent and shared definition of good teaching and be implemented by skilled evaluators. Teacher evaluation is essential to promote professional development and without a clear assessment of teachers' and administrators' understanding regarding the evaluation process, professional development for improved instruction and student achievement are absent or lacking clear objectives for goals (Danielson, 2011). Although evaluation systems had been in place for years, KEEP had been used in District G for 1.5 years at the time of the current study (director of human resources, personal communication, July 1, 2015). District G needs to determine teachers' and administrators' understanding of the differences among the teacher performance ratings and of the calculation of the final summative teacher evaluation rating. District G also needs to determine teachers' and administrators' perceptions of whether the district is using the KEEP model effectively.

Purpose of the Study

The first purpose of this study was to determine whether teachers and administrators understand the differences among the teacher performance ratings of highly effective, effective, developing, and ineffective. The second purpose of the study was to determine if teachers' understanding of the ratings was affected by school level and whether the teacher has been evaluated using KEEP and if administrators' understanding of the ratings was affected by school level. The third purpose of the study was to determine whether there was a difference between teachers' and administrators' understanding of the ratings and whether the difference was affected by school level.

The fourth purpose of the study was to determine teachers' and administrators' understanding of the calculation of the final summative teacher evaluation rating. The fifth purpose of the study was to determine if teachers' understanding of the calculation of the final summative teacher evaluation rating were affected by school level, whether the teacher has been evaluated using KEEP, and if administrators' understanding of the final summative teacher evaluation rating were affected by school level. The sixth purpose of the study was to determine whether there were differences between teachers' and administrators' understanding of final summative teacher evaluation rating and whether the differences were affected by school level. The seventh purpose of the study was to determine teachers' and administrators' perceptions of whether the district is using the KEEP model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions. The eighth purpose of the study was to determine if teachers' and administrators perceptions of the use of the KEEP model were affected by school level and by whether the teacher had been evaluated using KEEP.

Significance of the Study

The results of this study could determine whether teachers and administrators perceive they fully understand the calculation of the final summative rating and the six guidelines established within the Kansas evaluation process. In addition, the data gleaned from this research could determine whether school level (elementary [K-4], middle school [5-8], and high school [9-12]) affected teachers' and administrators' perceptions as well as the differences between both groups' perceptions. After perceptions are identified, a professional development plan could be created for the following year.

Other Kansas districts and the Kansas State Department of Education might utilize this research to modify the Kansas Educator Evaluation Protocol or the training provided to teachers and administrators who utilize the evaluation program. The findings of the study may be used to improve the evaluation process in District G and feedback to teachers provided by administrators, and adds to the body of research related to teacher evaluation.

Delimitations

According to Calabrese (2006), delimitations are the voluntary boundaries that are used to restrict the scope of the study. This study was delimited to administrators and teachers who were employed during the 2015-2016 school year by District G. Other districts in Kansas that use Kansas Educator Evaluation Protocol were not included in this population. The study was also delimited to administrators who evaluated classroom teachers. Lastly, two surveys (one for administrators and one for teachers) were created to determine the perceptions of teachers and administrators. The surveys were administered in November of 2015 and can be found in Appendix A and B.

Assumptions

"Assumptions are postulates, premises, and propositions that are accepted as operational for purposes of the research" (Lunenburg & Irby, 2008, p. 135). The researcher assumed that participants completed the survey independently, without seeking the support of or being influenced by peers. The second assumption was that all participants had some knowledge of the Kansas Educator Evaluation Protocol. The third assumption was that the list of teachers and administrators that used the Kansas Educator

Evaluation Protocol was current and accurate. The surveys utilized in the study accurately determined the perceptions of both teachers and administrators about KEEP.

Research Questions

According to Lunenburg and Irby (2008), the research questions guide the study, give it focus, and, serve as the "directional beam for the study" (p. 126). To guide this study, the following research questions were established.

RQ1. To what extent do teachers understand the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective?

RQ2. To what extent is teachers' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

RQ3. To what extent is teachers' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective affected by whether the teacher has been evaluated using KEEP?

RQ4. To what extent do administrators understand the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective?

RQ5. To what extent is administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12])?

RQ6. To what extent is there a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective?

RQ7. To what extent is the difference between teachers' and administrators' understanding of the differences between highly effective, effective, developing, and ineffective ratings of teacher performance affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

RQ8. To what extent do teachers understand the calculation of the final summative teacher evaluation rating?

RQ9. To what extent is teachers' understanding of the calculation of the final summative teacher evaluation rating affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

RQ10. To what extent is teachers' understanding of the calculation of the final summative teacher evaluation rating affected by whether the teacher has been evaluated using KEEP?

RQ11. To what extent do administrators understand the calculation of the final summative teacher evaluation rating?

RQ12. To what extent is administrators' understanding of the calculation of the final summative teacher evaluation rating affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

RQ13. To what extent is there a difference between teachers' and administrators' understanding of the calculation of the final summative teacher evaluation rating?

RQ14. To what extent is the difference between teachers' and administrators' understanding of the calculation of the final summative teacher evaluation rating affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

RQ15. To what extent do teachers perceive that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions?

RQ16. To what extent are teachers' perceptions that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

RQ17. To what extent are teachers' perceptions that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions affected by whether the teacher has been evaluated using KEEP?

RQ18. To what extent do administrators perceive that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions?

RQ19. To what extent are administrators' perceptions that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12])?

RQ20. To what extent is there a difference between teachers' and administrators' perceptions that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions?

RQ21. To what extent is the difference between teachers' and administrators' perceptions that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

Definition of Terms

According to Lunenburg and Irby (2008), key terms used throughout a researcher's dissertation should be defined. For this study, the following items are defined.

Developing. A developing rating is when an "educator sometimes exhibits an adequate level of performance on this component" (KSDE, 2014, p. 3).

Effective. An effective rating is when an "educator usually exhibits a more than adequate level of performance on this component" (KSDE, 2014, p. 3).

Evaluation. An evaluation is "a systematic determination of merit and significance of a person, program, organization, etc. using criteria against a set of standards" (KSDE, 2014, p. 127).

Highly Effective. A highly effective rating is when an "educator consistently exhibits a high level or performance on this component" (KSDE, 2014, p. 3).

Ineffective. An ineffective rating is when an "educator rarely exhibits an adequate level of performance on this component" (KSDE, 2014, p. 3).

Kansas Educator Evaluation Protocol (KEEP). KEEP is an educator evaluation system designed by KSDE in partnership with the Kansas education community. It was designed to "espouse support and acknowledgment of critical components of professional practice that ensures valid outcomes" (KSDE, 2014, p.1).

Performance criteria. Performance criteria are **"**levels of educator proficiency (highly effective, effective, developing, ineffective) that are used to evaluate performance, as specified in scoring guides such as descriptions or rubrics" (KSDE, 2014, p. 129).

Summative rating. A summative rating is an "overall rating of the level of performance based on the professional judgment of the evaluator considering all evidence and artifacts in the evaluation" (KSDE, 2014, p. 130).

Teacher performance level. A teacher performance level is a "descriptor of practice that is a valid measure supported by evidence and/or artifacts including measures clearly related to improving student growth" (KSDE, 2014, p. 3).

Overview of the Methodology

A descriptive non-experimental survey research design was used in this quantitative study. The study was conducted during the 2015-2016 school year. The participants in this study were teachers and administrators who used the Kansas Education Evaluation Protocol in a public school district in Johnson County, Kansas. Data was collected from both teachers and administrators via a survey created by the researcher. Twenty-one research questions guided this study. Each research question is

paired with an associated hypothesis and statistical analyses. Hypotheses were tested using one-sample *t* tests, two-sample *t* tests, one-factor analyses of variance (ANOVA), and two-factor ANOVAs.

Organization of the Study

This study contains five chapters. Chapter one included the introduction background, problem statement, significance of the study, delimitations, assumptions, research questions, the definition of terms, and an overview of the methods used during research. Included in chapter two are a review of the literature regarding the historical perspective on teacher evaluation, professional development associated with teacher evaluation, and perceptions of teachers and administrators on teacher evaluations. Chapter three is comprised of the research design; population and sample; sampling procedures; instrumentation, including measurement and reliability and validity; data collection procedures; data analysis and testing; and limitations of the study. Chapter four provides descriptive statistics, and the results of hypotheses testing. Lastly, chapter five includes a summary, findings related to the literature, and implications for action and suggested research to compose the conclusion of the study.

Chapter Two

Review of the Literature

Teacher evaluations have been heavily researched in recent years. Many school districts and states have accessed the available research to create evaluation tools due to No Child Left Behind (NCLB). NCLB marks the largest source of federal spending on elementary and secondary education. NCLB required states to ensure that all students were proficient in math and reading by 2014. Many states applied for a waiver from the NCLB requirements from the Department of Education. For the waiver to be accepted, the state needed to show that they had implemented a series of reforms to their academic standards, assessments, and evaluation systems for educators in K-12 schooling (NCLB). In 2012, the Obama Administration revealed the Race to the Top (RTT) program (White House, 2015). Incentives were provided to states willing to initiate new programs or procedures to support teaching and learning (White House, 2015). The reform includes continued development of rigorous standards and assessments, the inclusion of data systems for progress monitoring of students, professional development for teachers and administrators, and increased resources for interventions for the lowest-performing students and schools (White House, 2015).

Along with the four objectives of the reform, RTT has raised standards and aligned curriculum to the goal of college and career readiness. Prior to both RTT and NCLB, the publication of *A Nation at Risk* in 1983 created "the push for teacher quality and has developed the modern school reform movement" (Danielson, 2001, p. 2). Each of these national school reform movements has taken part in shaping the current perspective of teachers and administrators on teacher evaluations. Chapter two contains a

review of research and literature related to the historical perspective on teacher evaluations, professional development associated with teacher evaluations, and a review of the recent research conduct over teachers' and administrators' perceptions of supervision and evaluation.

Historical Perspective on Teacher Evaluation

Due to political and social changes, the process of teacher evaluation has evolved since the early 18th century, and many notable educational leaders have influenced the evolution of the evaluation processes utilized in schools (Marzano, Frontier, & Livingston, 2011). Throughout history, many individuals have been responsible for the evaluation of teachers within a community (Marzano et al., 2011). The responsibility for evaluations has most recently been communicated in a structured process conducted by administrators (Marzano et al., 2011). According to Tracy (1995), the historical perspective on teacher evaluation can be outlined in seven distinctive phases: community accountability phase, professionalization phase, scientific phase, human relations phase, the return of scientific phase, the return of human relations phase, and human development phase.

Community accountability phase. In the first phase, communities were responsible for evaluating the effectiveness of teachers. In the 1700s, towns in the United States turned to existing local governments and clergy to hire and evaluate teachers (Marzano, Frontier, & Livingston, 2011). Committees created by the community leadership would visit the schools and ensure that appropriate curricular content was being taught. The committee members were responsible for speaking directly with the teacher when objectives or expectations were not met (Tracy, 1995). The community

accountability phase also gave committee members the power to establish criteria for effective instruction. Community members now had the power to hire and fire teachers (Burke & Krey, 2005). The evaluations were done to inspect their teaching, moral standards, and ethical choices within the community.

Professionalization phase. The next phase, which extended through most of the 1800s, focused on professionalization (Tracy, 1995). The professionalization phase added supervisors, who were thought to be experienced and valuable teachers within a school community. Supervisors were responsible for evaluating and aligning community morals and ethics in the classrooms. This "principal" supervisor ultimately grew into the role of the modern day principal (Marzano et al., 2011). The principal is frequently the data gatherer and decision maker for annual teacher evaluation. The principal's actions consisted of gathering data through observation, some form of clinical supervision, and anecdotal records. The decision is then made for dismissal, remediation, or retention.

Scientific phase. The scientific phase was from the 1900s through the 1920s (Tracy, 1995). The effective evaluation of a teacher began to emerge, and supervision became a formal activity. Cubberley (1929) applied the management guidelines of a factory to the management principles of a school. Cubberley (1929) stated, "Our schools are, in a sense, factories in which the raw products (children) are to be shaped and fashioned into products to meet the various demands of life" (p. 338). The earliest documented evaluation instruments emerged in 1915, making evaluations more formalized and structured (Medley, 1997). Evaluations consisted of measuring control of the classroom, accountability, and efficiency of teaching (Tracy, 1995). Guiding principles were outlined within the evaluation instruments; school principals could

analyze data to support effective teachers and schools. Cubberley provided specific examples of applying the scientific approach to observations and feedback a principal could provide a teacher. As a result, effective teaching traits were identified through evaluation instruments that provided a firm base on which to judge the quality of instruction. Tracy (1995) proclaimed, "Rather than simply understanding the mores of the community, the supervisor now needed to have subject area knowledge and teaching skills" (p. 323). Supervisors began to conduct direct classroom observations and data gathering. An observation checklist was also a tool that evaluators used during the scientific phase.

Human relations phase. Human relations was the focus of the fourth phase, occurring in the 1930s through the early 1940s. Supervisors shifted to view teachers as individuals within a school organization. The supervisors' main goal then became to support the professional development of the teacher (Tracy, 1995). Administrators worked to develop relationships with their team members and began to view their role as a mentor. Principals became active in the development of curriculum, staff development, and the relationship the school had with the community (Glanz, 1991). Principals were no longer evaluators but became team members required to improve instruction throughout the school. Evaluations consisted of measuring a teacher's contribution to curriculum and collaboration. Administrators concentrated on building positive relationships and not conducting direct classroom observation, as direct observations were thought to upset the relationship between the teacher and administrator (Tracy, 1995).

Return to the scientific phase. The fifth phase, occurring in the late 1940s to the early 1960s, was the return of the scientific phase. This phase focused on the goal of aligning objectives within an evaluation to desirable teacher traits. Supervisors provided feedback on teacher characteristics such as data collection, relating to people, and collaboration with teammates. The superintendent selected supervisors, who conducted the evaluations. Superintendents based their selection of supervisors on successful teaching experience. Educators and supervisors felt they needed to focus on balancing the technical aspect of teaching with the autonomy of the teacher when applying learned skills (Tracy, 1995). Evaluation models began to become more collaborative with school teams. The model consisted of the supervisor and teacher analyzing the teacher's performance together. A sustained cycle of assistance from the supervisor was a necessary element of improvement in teacher behavior. Pre- and post-observation conferences also took place within a cycle, allowing the teacher and supervisor to solve classroom problems (Tracy, 1995).

Return to the human relations phase. The sixth phase, occurring in the mid-late 1960s through the mid-1980s, focused on effective instructional models (Danielson & McGreal, 2000). Various observation instruments were developed for use when conducting classroom observations. The connection between effective teaching strategies was linked to student achievement. Studies were then conducted to identify appropriate teaching behaviors that were directly linked to the student attainment of basic skills (Danielson & McGreal, 2000).

Hunter became a major influence during this phase, developing the lesson plan decision-making model (Marzano et al., 2011). Hunter's model helped teachers to

describe their lesson and supervisors to determine the effectiveness of the lesson (Hunter, 1980). The decision-making model included anticipatory set, objective and purpose, input, modeling, checking for understanding, guided practice, and independent practice. Along with the decision-making model, Hunter also began to contribute many other ideas to the process of supervision. She created professional development to articulate a common language of instruction and identified a variety of purposes for supervisory conferences (Marzano et al., 2011). Script-taking or observations were important components of Hunter's process of supervision. During script-taking, administrators would document behaviors of effective teaching and categorize the behaviors into those that "promoted learning; those that used precious time and energy, yet contributed nothing to learning; and those that, unintentionally, actually interfered with learning" (Hunter, 1980, p. 409). After script-taking, administrators would meet with the teacher to discuss the notes from the observations (Marzano et al., 2011).

McGreal created a range of supervisory categories for teachers linked to their teaching experience. The supervisory categories for teachers ranged from intensive developmental supervision for novice teachers or teachers in need of growth plans to self-directed professional development for experienced staff (Marzano et al., 2011). McGreal (1983) recommended that teachers be placed in one of two categories, an intensive evaluation program for personnel decisions or in a standard evaluation program designed for continual professional development. McGreal (1983) outlined the importance of goal setting and planning within the teacher evaluation process. McGreal believed there were two ways to increase the reliability of classroom observations. First, goals need to be made, allowing the principal to focus the objectives of the observation. McGreal stated,

"The goals become the core of the evaluation/supervision process" (1983, p. 305). Secondly, a pre-conference needs to be conducted to provide an overview for the teacher on what the principal is going to observe (McGreal, 1983).

Also during the 1980s, the RAND group conducted a study to determine the types of evaluations practices that were occurring in school districts across the nation. The results of the study indicated that many of the evaluations protocols in place were didactic and not specific enough to enhance pedagogical development (Wise, Darling-Hammond, McLaughlin, & Bernstein, 1984). Teachers wanted a more standardized and formulated process for ratings. "In their view, narrative evaluation provided insufficient information about the standards and criteria against which teachers were evaluated and resulted in inconsistent ratings among schools" (Wise et al., 1984, p. 16). Teachers were also resistant to the feedback provided by principals due to the lack of consistent evaluation protocols throughout a school building (Wise et al., 1984).

Human development phase. Due to educational reform, the most significant paradigm shift in education is presently occurring. Teacher evaluation now focuses on teacher professional development and accountability for student learning or achievement scores. New evaluation models focus on professional practice based on professional teaching standards. Darling-Hammond (2012) stated, "These protocols generally provide indicators of teaching effectiveness associated with a set of professional standards that are concrete enough to guide observations and feedback to teachers—the standards describe practices shown by research to be associated with student learning" (p. 16). Researchers have learned that teacher effectiveness was vital in supporting teachers in developing effective student learning outcomes (Darling-Hammond, 1999) and that

teacher supervision was essential for increasing teacher practice and student learning successfully (Sullivan & Glanzer, 2000). Danielson and McGreal (2000) describe six deficiencies of evaluation systems used in the 1990s: The evaluation system had outdated and limited evaluative criteria, the perceptions of what good teaching looks like were varied, the rating scales lacked validity, differentiation between new and experienced teachers was not addressed, and communication was usually evaluator-driven, and many administrators lacked expertise in content areas.

In 1996, Danielson published her work on supervision and evaluation. Her research was conducted with the Educational Testing Service and the measurement of the competence of new teachers (Marzano et al., 2011). The Danielson model contains four different domains: planning and preparation, the classroom environment, instruction, and professional responsibilities (Danielson, 2008). Danielson's framework was created to bring light to the complexity of teaching, create a language for professional growth, and provide a structure for self-assessment and reflection for educators (Marzano et al., 2011). The comprehensive framework also includes four levels of performance and 76 elements of quality teaching characteristics (Danielson, 2008). She stated, "Educators have found that when their discussions are organized around a clear definition of good teaching, the conversations themselves are more productive and focused" (p. vii). Danielson's framework for teaching also contains a self-assessment. Teachers locate examples of their teaching that supports the different components of the framework; this ensures that they see the connection to their teaching within the framework (Danielson, 2008).

The Danielson teacher evaluation model was a significant enhancement for evaluating teachers. Danielson's model consists of three different tracks. Track One: an administrator is encouraged to spend more time mentoring new teachers to focus on the evaluation practice and ensure that accurate statements are being made regarding the summative evaluations (Danielson & McGreal, 2000). Track Two: allocates less time to experienced teachers who already have established a record of highly effective instruction. Administrators are encouraged to foster professional growth opportunities to promote continued skills through activities such as professional learning communities, action research, curriculum development, peer coaching, professional portfolios, and study groups (Danielson & McGreal, 2000). Track Three: focuses on the needs of developing or non-effective teachers by providing more rigorous assistance and clear standards for development for these teachers (Danielson & McGreal, 2000). Danielson's model recognizes various characteristics and components that support exemplary teaching. Supervisors are also encouraged to collect data on teacher planning, communication with stakeholders, and participation in professional development. All the evidence of effective teaching is collected in other interactions beyond the classroom observation. One of the significant characteristics of the Danielson model is that it provides constructive feedback to ensure teacher growth and the evaluation of teachers is differentiated based on the starting point of the teacher (teacher's experience, goals, and development). Teachers are given targeted feedback with specific areas of improvement for their goals.

Weisberg, Sexton, Vlulhern, and Keeling (2009) conducted a study in which 12 school districts in Arkansas, Colorado, Ohio, and Illinois participated. The Widget Effect

researchers pointed out that school districts and states assume effectiveness is the same for each teacher and that 99% of teachers were usually categorized as good or great. Administrators also neglected to add comments to 73% of the evaluations. Half of the districts studied had not dismissed an educator for low ratings in the past five years from when the study took place (Weisberg et al., 2009). The results of the research of the Widget Effect showed the lack of training an administrator obtained to complete evaluations and rate the performance of teachers. As Weisberg et al. (2009) stated, these administrators were unsure whether the "system drives the culture, or the culture the system" (p. 6).

The Marzano Teacher Evaluation Model was created in 2011 and was based on a number of his related works (Marzano, 2011). The Marzano Teacher Evaluation Model is composed of four domains (classroom strategies and behaviors, preparing and planning, reflecting on teaching, and collegiality and professionalism) and 60 elements (Marzano, 2011). Within the four domains and 60 elements of effective teaching, there are strategies for principals to guide the development of teachers over a year (Marzano, 2011). Some states have used or adapted Marzano's model as the basis for teacher feedback and evaluation.

Professional Development Associated with Teacher Evaluations

Danielson and McGreal (2000) stated, "The challenge confronting designers of an evaluation system is to (1) encourage professional learning and, at the same time, (2) ensure the quality of teaching" (p. 21). Professional development is the ongoing learning that leaders within a district provide their teachers. "Research supported that educators had little guidance in the development of their professional skills and educator goals were

not stated by their leaders in ways that they could understand and apply to their growth" (Johnson, Dupuis, Musial, Hall, & Gollnick, 1995). Professional development is most effective when it is completed in a collaborative effort. Dyer and Renn (2010) outlined four principles of adult learning: "Adults need to know why they need to learn something, adults need to learn experientially, adults approach learning as problem solving, adults learn best when the topic is of immediate value" (p. 4). Danielson and McGreal (2000) stated, "The principles of adult learning show that when people use self-assessment and self-directed inquiry in professional development, they are more likely to sustain their learning, in more disciplined ways, than when outsiders impose professional development requirements" (p. 25). Within an effective teacher evaluation, teachers should engage in reflection, self-assessment, and collaboration to ensure adequate growth in professional development (Danielson & McGreal, 2000, p. 24).

Many districts have aligned their teacher evaluation systems to their school professional development plans. Danielson and McGreal (2000) stated that school-wide professional development plans allow school teams to "establish professional development plans within the evaluation system over multiple years" (p. 18). The principal serves at the facilitator of the professional development plan. Danielson (2008) stated, "Administrators bear a certain responsibility for establishing and maintaining a culture of professional inquiry with a school" (p. 17). Through the development of the school professional development plan, school teams have the opportunity to complete complex activities that can range over multiple years (Danielson & McGreal, 2000). Once the team reaches closure on the plan, they need to establish a new professional

development plan, “there should be no time off from growing professionally” (Danielson & McGreal, 2000, p. 19).

Evaluation systems have also led to the development of individual learning plans to ensure ongoing professional development for the teacher. Evaluators and teachers use the “information gathered from observation to create an individual professional learning plan that lays out what learning opportunities the teacher may have to advance her practice, with individual goals tied to school and district goals” (Frank, 2013, p. 4). Within the individual learning plan are goals for the teacher to accomplish within a certain period. The evaluator and teacher can both formulate professional development that supports the goals together. Culbertson (2012) stated that "When evaluation systems are focused on improving practice in addition to measuring performance, they yield powerful results" (p. 14).

When teachers are demonstrating that they are not meeting the goals of their individual or school professional development plan, the principal may choose to begin an assistance plan. A school district must determine when an assistance plan or plan of improvement is to be utilized (Danielson, 2008). Danielson (2008) stated that typically, “Unsatisfactory level on any component of the framework for teaching results in an assistance plan to address that issue” (p. 49). The administrator, teacher, mentor, and instructional coach can develop the assistance plan together. The assistance plan can be in place for a few weeks or years. The duration of the plan is dependent on the growth exhibited by the teacher and noted by the principal. Danielson (2008) stated, “The key to promoting professional learning lies in the procedures adopted and the culture within which the framework is used” (p. 24). Teachers should feel comfortable to direct their

own professional learning and conduct self-directed inquiry. When creating evaluation systems, ensuring they contain reflection, self-assessment, and collaboration are the best way to ensure that teacher is as active as possible in their learning. When teachers reflect on their teaching practices, they are able to identify areas of teaching that need to be strengthened (Danielson, 2007). The act of reflecting can be completed throughout the whole evaluation process. Reflection should be done privately by the teacher or with the aid of a mentor or peer coach (Danielson, 2007).

During evaluations, supervisors can support appropriate reflection for a teacher by structuring the conversation during the post observation. Teachers should be aware that the lesson does not have to be perfect, but they should be able to identify the strengths and weaknesses of the lesson. Danielson (2007) explained, "The key to being an accomplished teacher is acquiring the skill to continually improve one's practice, and important vehicle for this is reflection and conversation" (p. 169).

Reflection is intertwined with self-assessment; a teacher self-assesses and reflects throughout the practice of teaching. Evaluations should include a portion devoted to self-assessment. The self-assessment portion of the evaluation is the starting point for the communication between the observer and the teacher. Self-assessment can be completed in a self-rubric or analysis of data. Danielson and McGreal (2000) believed, "When people select their own "problem" to be solved, their own project to pursue, they devote greater energy to it than if someone has chosen the issue" (p. 25). Teachers are the guides to their professional growth throughout their career. Self-assessment can also be completed in collaboration with the evaluator, especially with novice teachers (Danielson & McGreal, 2000).

Collaboration within an evaluation system allows for a balanced and more accurate interpretation of the practice (Danielson & McGreal, 2000). In many districts, collaboration is conducted by the teacher with the mentor, teaching team, or supervisor. Collaboration and mentorship have proven to be effective in enhancing the quality of teachers (Danielson, 2002). Teams and partners within a school team require teachers to be active participants, which creates a culture that encourages reflection, constant peer learning, and job satisfaction through a feeling of belonging (Danielson, 2002).

Danielson (2008) stated, "The key to promoting professional learning lies in the procedures adopted and the culture within which the framework is used" (p. 24). The teacher should feel comfortable to direct his or her professional learning and conduct self-directed inquiry. When creating evaluation systems, ensuring they contain reflection, self-assessment, and collaboration are the best ways to ensure that teacher is as active as possible in their learning.

In 2003, the Iowa State Legislature adopted the Iowa Teaching Standard and a Model Framework for Teacher Evaluation, which was influenced by Danielson's Framework for Teaching. In 2011, Huckstadt conducted a study comprised of two distinct components: a quantitative descriptive analysis of perceptions of teachers and administrators regarding the standards-based teacher evaluation system in Iowa, and a comparison of teacher and administrator survey responses. The study Huckstadt (2011) concluded that teacher evaluation should be focused on the formative aspects of evaluation, using staff-directed activities for promoting professional development, especially development focused on improving student achievement as determined by district achievement goals.

An additional study conducted in Tennessee reviewed the effects of teacher collaboration on student achievement or teacher growth. Vanderbilt University's Peabody College and the Tennessee Department of Education are using the U.S. Department of Education's Institute for Education Sciences grant to find ways that evaluations can build stronger school culture and growth among their teachers through reflection or collaboration (Gonzales, 2015). The focus of the study was on teacher partnerships between each other and their growth due to the Instructional Partnership Initiative (Gonzales, 2015). The goal of the partnerships is to provide professional development to teachers at no cost to the district (Gonzales, 2015). The initiative's effectiveness was tested in 1,453 schools in Tennessee (Gonzales, 2015).

Teachers' and Administrators' Perceptions of Teacher Evaluation

As new evaluation systems have become widespread in school districts, teachers' and administrators' attitudes regarding the evaluation system have played a critical role in the success of an evaluation system. Larsen (2004) explains that teachers tend to view the evaluation process as threatening and stressful. Danielson and McGreal (2000) noted that teachers are reluctant to trust the ratings given in an evaluation: "Although all teachers want the highest rating and believe that their careers will be damaged if they don't receive it, they believe that administrators reserve the highest ratings for their friends or protégés" (p. 5). Morelock (2008) determined in her study that effective teacher evaluations begin with a common understanding of a quality teacher. Also, administrators need to have an awareness of instructional practices. Administrators must view instruction as their top responsibility above other administrative responsibilities, consistently looking for ways to provide professional development for their team and

improve learning for every student. Visits to classrooms must be constant, and collaboration needs to be embedded into the weekly schedule for teachers (Morelock, 2008).

In 1993, the National Center for Educational Statistics (NCES) conducted a study on teachers' perceptions of teacher evaluations. This study was the initial research conducted with teachers that taught kindergarten through sixth grade and their perceptions of performance evaluations. The results of the study indicated that 89% of the teachers felt their evaluations were an accurate indicator of their teaching performance, 63% stated they could design a plan for their professional development based on their last performance evaluation, and 74% felt that their last evaluation had positively influenced their teaching skills (NCES, 1994). All participants in the NCES reported that the evaluation encouraged the effectiveness of their teaching (NCES, 1994).

In 1994, Rindler surveyed 435 teachers involved in the University and College Intensive English Programs. Teachers reported there were several factors that impacted their professional growth from the results of an evaluation. These factors included the rationale for suggestions provided in the evaluation, the level of trust the teacher had for the evaluator, modeling that took place by the evaluator, the specificity noted in suggestions made, the amount of credible information contained in the feedback, the diligence done during the evaluation, the evaluation focus on standards or goals created by the teacher, the true role of the evaluation, and the teacher's prior experience with supervisors and evaluation models (Rindler, 1994).

In 2003, Zimmerman and Deckert-Pelton surveyed 86 teachers from various districts in Florida. The survey had written responses and open-ended questions. The

focus of the study was on educators' general perceptions of teacher evaluation. Some of the teachers believed that effective evaluations included open conversations about their lessons and the opportunity to discuss the feedback they received. Overall, the Zimmerman and Deckert-Pelton (2003) study concluded that teachers' perceptions of an effective evaluation model depended upon the instructional leadership of the principal.

In 2008 and 2010, Breedlove (2011) analyzed the Teacher Working Conditions Survey (TWC) to determine if the perceptions of the teachers on the evaluation process had changed after the revisions of the North Carolina's evaluation process. The revisions done prior to the survey included establishing clear standards for the evaluation process and using a rubric to assess the standards. Like the KEEP evaluation, self-assessment was added to the process and the collection of artifacts throughout the cycle. Breedlove collected over 105,600 responses to the survey from 2008-2010 (Breedlove, 2011). The results of the study indicated that a majority of teachers felt that the revisions to the evaluation process were positive modifications, although alterations were still needed such as "consistent implementation, further guidance on goal setting and the development of professional development plans, additional observations and a focus on student performance and outcomes instead of primarily focusing on 'teacher actions'"(Breedlove, 2011, p. 145).

Pizzi published a study in 2009 that highlighted teacher's perceptions of the "effectiveness of standards-based evaluation systems as a means of improving instruction and student achievement" (p. 3). The study was conducted at a large urban high school located in the northeastern United States. Pizzi found that a majority of the teachers were not familiar with the standards in the district evaluation tool. Teachers also believed that

the conferences with administrators were the most significant resources in helping them improve their instruction within the classroom. In addition, the existing evaluation tool lacked a connection to annual goals or professional development plans and multiple sources of data were not included in the summative rating for educators. Overall, Pizzi found that there was not a connection between standards-based performance evaluation tools and the improvement of student achievement (Pizzi, 2009).

Princess Towe (2012) sought to find the perceptions of administrators, language arts teachers, and math teachers in four high schools in a New Jersey school district. The district implemented a standards-based evaluation system in 2003 (Towe, 2012). The framework was modeled after Danielson's 1996 publication, Enhancing Professional Practice: A Framework for Teaching (Towe, 2012). The focus of the study was to seek a correlation between the implemented teacher evaluation process in the New Jersey school district and the administrators' roles in teacher evaluation process. Development of professional growth opportunities and teacher practice were examined. The results concluded that teachers and administrators within the high schools perceive the current evaluation system being implemented to have little effect on improving the effectiveness of teachers (Towe, 2012). The summative evaluation has the most effect on professional development goals. Towe (2012) suggest three goals for the district that include implementing measures to ensure fidelity of the evaluation process within buildings, examine the impact of the multiple data sources (artifacts, self-evaluation, and peer evaluation) on professional growth, and research the district's specific needs for teachers and administrators to promote a growth-oriented teacher evaluation system.

In 2012, Ruffini, Makkonen, Tejwani, and Diaz (2014) conducted a study in Arizona school districts after their initial year with the new Arizona Framework for Measuring Educator Effectiveness. The new evaluation system included three components: teaching observations; surveys of students, parents, and peer teachers; and measures of student academic progress (Ruffini et al., 2014). The findings from the study indicated that 39% of teachers from the piloted districts agreed that their final summative classification was accurate, 32% believed their summative rating was not accurate, and 30% were indifferent regarding the final rating (Ruffini et al., 2014). Teachers also reported that they were more reflective throughout the evaluation process and were open to measures involving student performance and stakeholder surveys, but only if "they perceived the metrics to be fairly and consistently applied" (Ruffini et al., 2014, p. 3).

Bonavitacola (2014) investigated teachers' perceptions of the impact of the Mid-continent Research for Education and Learning (McREL) Teacher Evaluation System on their professional growth completed through a qualitative study. Fifteen teachers in an elementary school in New Jersey participated in the one-year implementation of the new standards-based McREL Teacher Evaluation model. After analyzing documentation, survey data, and interviews conducted at the elementary school, it was perceived that different components of the evaluation process affected teaching practices in the classroom (Bonavitacola, 2014). Most importantly, Bonavitacola (2014) concluded,

> The culture of the school site had also been impacted by the importance placed on collaboration to improve teaching and learning, as indicated by the 91.4% of the teachers who collaborated on a regular basis on instruction matters. As stated in

> the interviews by those who participated in an alternative option, the partnership option was seen to be a strong method of supporting teachers' professional development because of its collaborative and reflective nature. (p. 117)

Professional development and collaboration have shown to have positive effects on improving teachers' practice.

Administrators have also developed perceptions regarding teacher evaluation systems. Some administrators complain about how time-consuming the evaluations process is with all the other daily responsibilities they have in leading schools. Other administrators consider teacher evaluations are the worst part of their job (Hopkins, 2001). Research has shown that some administrators believe evaluations are used for the purpose of either retaining quality educators or dismissing the unsatisfactory teachers within a school (Sutton, 2008). Administrators believe there are several barriers to conducting effective evaluations. Administrators struggle to find access to data to fit within the standards of an evaluation process. Administrators also can lack an understanding of statistical models used in the evaluation system (Summerville, 2014). Last, the district provided little to minimal training in the use of analyzing available data for the purpose of teacher evaluation (Summerville, 2014).

Marshall (2005) published ten reasons why the "Conventional supervision and evaluation process is not an effective strategy for improving teaching and learning" (p. 728). Marshall first stated a belief that principals only evaluate a small portion of teaching. Teachers are not monitored for a high percentage of the year; they are trusted to be professional and competent in their area. The second reason that conventional supervision and evaluation is not effective for improving teaching and learning is that

evaluations of only a portion of the year do not heavily influence teachers to modify their practice. "Many school districts try to compensate for how little time principals spend in individual classrooms by requiring extremely thorough evaluations of lessons that are formally observed" (Marshall, 2005, p. 728). Evaluations are rarely reviewed by other leaders within the district, except in extremely rare cases when a supervisor rates a teacher as unsatisfactory (Marshall, 2005).

Third, principals evaluate "atypical" lessons (Marshall, 2005). Teachers are able to prepare the lesson with advanced notice and to have a "top-level authority" in the classroom reduces discipline problems. Although both these factors can work to benefit a teacher's overall rating, negative factors can also take place. Some teachers get nervous and end up failing their evaluation for one year in a thirty-minute period.

Additionally, the fourth reason is that isolated lessons give an incomplete picture of instruction (Marshall, 2005). Observing one lesson does not accurately display the continuum of learning. Supervisors hardly view unit plans or assessment results in evaluations. Observing lessons over a three or six week period would give a better estimate over which curriculum are aligned.

The fifth reason highlights that evaluation hardly focuses on student learning. Many school districts have had teacher unions push back on the thought of tracking student success. "Even the most fervent advocates don't think it's fair to use them to evaluate a teacher after only a year of instruction" (Marshall, 2005, p. 730). Sixth, high-stakes evaluation is not a culture of learning for adults. Teachers consistently fear that when they are evaluated their jobs are on the line. Many times the principal conducts the feedback and own the conversation; the teacher has little to contribute because it is an

uncomfortable setting to be authentic. Seventh, supervision reinforces the separation of teachers and administrators. Evaluations are conducted in private meetings and confidential documents; they are never used to encourage collaboration among a team of teachers. Eighth, evaluation instruments often get in the way of a supervisor simply mentoring a teacher. Evaluation tools "are rarely conducive to fostering an honest, open, and pedagogically sophisticated dialogue between principals and teachers" (Marshall, 2005, p. 731). The ninth reason is that evaluations fail to give teachers constructive feedback. Principals give "satisfactory" ratings and neglect to articulate clear standings on performance standards. The last reason, most principals find it impossible to support good supervision and respectable teacher evaluations. Discipline and operating duties are extremely demanding that the teacher evaluation often fades from principals' duties until contractual deadlines force them to complete the evaluations (Marshall, 2005).

Various notable studies were conducted throughout 2006 to 2014 regarding administrators' perceptions of teacher evaluations. Kersten and Israel (2005) conducted research with K–8 school building administrators in one section of one county in Illinois. An 18-item self-administered questionnaire was given to participants (Kersten & Israel, 2005). Overall, Kersten and Israel (2005) found that administrators view current evaluation practices as not as effective due to the limited time administrators can spend on the process and the unclear goals or feedback provided by district procedures.

A qualitative case study of teacher evaluation and supervision at a high-achieving urban elementary school was conducted by Lansman (2006). The purpose of the study was to determine the impact of teacher evaluation on teacher performance. The sample group in Lansman's (2006) study consisted of teachers, a site principal, and a district

administrator within a Los Angeles county urban elementary school. Lansman (2006) sought to find key elements of the school's success of the evaluation process on teacher evaluation, teacher effectiveness, and student achievement. Data collected by researcher consisted of district policy documents, conducting classroom observations, collecting confidential surveys from 19 classroom teachers, and conducting one-on-one interviews with ten classroom teachers, the site principal, and the district assistant superintendent of human resources (Lansman, 2006). The analysis of the data collected concluded that within the successful urban elementary school that strong leadership, collaboration, and accountability impact the effectiveness of the teachers and the impact the evaluation process have on positively affecting teacher practice and students achievement (Lansman, 2006).

Additionally, Garth-Young (2007) found in his study that middle school and junior high principals in Illinois believed "time constraints" and "inadequate instrumentation" for evaluating teachers were hindrances for administrators (p. 102). Garth-Young (2007) argued that "Quality evaluations may be possible if the amount of time to conduct evaluations and the number of teachers to be evaluated were within reasonable parameters" (p. 124).

In 2009, Doherty surveyed fourteen administrators in Massachusetts. These administrators believed that improvements could be made to the evaluations system used by "differentiating the teacher evaluation system, reducing the amount of paperwork in the process, increasing the number of informal observations and walkthroughs, developing differentiated rubrics for different teaching positions, and using multiple sources of data" (Doherty, 2009, p. 4).

Administrators are responsible for ensuring the evaluations are completed, and the process is followed with fidelity. Shana Henry Barton (2010) investigated K-12 principals' perceptions of teacher evaluation in an urban northern California school district. Fifty-two of the 70 surveys distributed to school principals were returned (Barton, 2010). Barton's (2010) survey contained closed and open-ended responses. Barton's findings illustrate that administrators find formative and/or summative evaluations most effective for novice teachers (2010). Tenured teachers responded most to formative evaluations (Barton, 2010). Not surprising, participants shared that the evaluation process within their district was time consuming and lacked a clear purpose for professional development or staffing decisions (Barton, 2010). Inconsistency of implementation by principals and lack of support from the district teacher union were all notable obstacles mentioned by the sample group were also concluded by Barton (2010).

Both quantitative and qualitative approaches were used sequentially in a mixed methods study conducted by Fox- Norwitz (2013). The researcher sought factors to support the implementation and sustainability of the Rhode Island Model of the Educator Evaluation System (Fox-Norwitz, 2013).

> Two main research questions were addressed: What is the relationship of the leadership dimensions of vision, support, structure, and trust on charter school teacher attitudes toward the implementation of the Rhode Island Model of the Educator Evaluation System? What do teachers perceive as the factors that contribute to teachers' attitudes toward the implementation of the Rhode Island Model of the Educator Evaluation System in relation to leadership? (Fox-Norwitz, 2013, p. xi)

Survey participants (Rhode Island grade 6 - 12 Charter School teachers) reported that structure, support, and trust affect their attitude towards the implementation of the Rhode Island Model of the Educator Evaluation System (Fox-Norwitz, 2013). A focus group was also conducted, which resulted in transparency and alignment of vision be reported as factors that affect teachers' attitudes toward the implementation of the Rhode Island Model of the Educator Evaluation System (Fox-Norwitz, 2013).

In a dissertation completed by Nixon, Packard, and Dam (2013), school administrators from Alabama, Georgia, North Carolina, and South Carolina completed a survey that "investigated the relationship between teacher dispositions, subject content knowledge, pedagogical content knowledge, and reasons that school principals recommend non-renewal of probationary teachers' contracts" (Nixon et al., 2013, p. 1). Administrators reported that they saw a lack of pedagogical content knowledge from unsuccessful teachers. Instructional skills were also noted as a priority when deciding whether to terminate a teacher each year. Important dispositions that were sought in teachers were integrity, dependability, and honesty (Nixon et al., 2013). Overall, the study concluded that administrators' views sought dispositions as personality characteristics rather than as competencies for a teacher.

Glowacki completed a study in 2013 on the perceptions of Illinois elementary school administrators on their existing teacher evaluation process and the practice of evaluating special education teachers. Another component of the study was for administrators to self-assess their efficacy in evaluating special education teachers. Three hundred thirty participants completed the online survey, of which 83.9 % reported that the current evaluation system used in the district did not differentiate between the

professional responsibilities of general and special education teachers (Glowacki, 2013). Lastly, elementary administrators selected various ways the district evaluation process could be modified to fulfill the correct reporting of job responsibilities of teachers. One of the primary changes being specific job responsibilities of special education teachers added to an evaluation (Glowacki, 2013).

In 2014, Summerville conducted a study at Vanderbilt University. Researchers at Vanderbilt University found that many principals did not use data to make informed decisions for evaluations; 84% trusted the teacher observation to help make these decisions (Summerville, 2014). The results exposed the perceived barriers for principals to access data and a lack of understanding of the use of data (Summerville, 2014). Districts needed to create training for principals on how to use and find data to understand the information for important personnel decisions.

In 2013-2014, the Ohio Teacher Evaluation System (OTES) was implemented (Dolph & Kowalski, 2015). Dolph and Kowalski (2015) examined Ohio principals' dispositions at the end of the school year after the principals had evaluated teachers under OTES for the first cycle. Eighty-nine principals employed in public elementary and secondary schools were included in the study population (Dolph & Kowalski, 2015). A panel of experts, current professors and administrators, developed the survey (Dolph & Kowalski, 2015). A return rate of 56% paper administered surveys were returned and analyzed between May and June of 2014 (Dolph & Kowalski, 2015). The results revealed from Dolph and Kowalski's (2015) study after the first year of implementation included not a sufficient amount of time to implement the evaluation process with fidelity. Administrators were also concerned that a teacher's performance was to be

based heavily on student growth within one year (Dolph & Kowalski, 2015). Additionally, administrators included in the study noted the difficulty of finding time to assist the teachers within their building on their annual improvement plans (Dolph & Kowalski, 2015). Most alarming, the bulk of the participants did not think that OTES would produce positive outcomes in several critical areas including (a) overall school improvement, (b) principal-teacher relationships, and (c) the validity and reliability of performance evaluations (Dolph & Kowalski, 2015).

Summary

This chapter reviewed the literature related to teacher evaluation systems, including an overview of the historical perspective and influencers. The professional development done in response to evaluations was also reviewed through the outline of reflection, self-assessment, and collaboration. Lastly, perceptions of teachers and administrators were explored, as attitudes influence the success of an evaluation system. In chapter three, the methodology and research design are presented.

Chapter Three

Methods

This research study was conducted to determine teachers' and administrators' understanding of the differences among the teacher performance ratings and of the calculation of the final summative teacher evaluation rating and to determine teachers' and administrators' perceptions of whether the district is using the KEEP model effectively. Additionally, the study was conducted to determine whether there were differences between the teachers' and administrators' perceptions. Finally, the effects of teacher and administrator levels and whether the teacher had been evaluated on the teachers' and administrators' perceptions and the differences between the perceptions were determined. Chapter three includes the research design; population and sample; sampling procedures; instrumentation, including measurement, reliability, and validity; data collection procedures; research questions and corresponding hypotheses and data analysis; and limitations of the study.

Research Design

This quantitative study was descriptive non-experimental in its research design. Using survey research, teachers and administrators who used the Kansas Education Evaluation Protocol in District G were administered a survey to determine their understanding of the evaluation process and their perceptions of the implementation of the KEEP model within District G. Dependent variables included the understanding of the teacher performance ratings, understanding of the final summative evaluation rating, and perceptions of district's use of the KEEP model. The school level of the teachers and

administrators and whether the teachers were evaluated using KEEP served as independent variables.

Population and Sample

The population chosen for this study was teachers and administrators utilizing the KEEP in Kansas. The sample was comprised of teachers and administrators employed by District G who completed the survey. Portions of teachers and administrators within District G have been using the KEEP evaluation tool for two years. Other participants have only utilized the KEEP for one year or have not used the KEEP. All teachers and administrators within the sample had been trained on the KEEP for two consecutive years.

Sampling Procedures

This study used a purposive sampling of teachers and administrators listed in the 2015-2016 District G Directory (human resources secretary, personal communication, November 19, 2015) and who completed the survey in its entirety. Administrators chosen for the study participated in the evaluation of teachers using the KEEP model and received professional development about the KEEP model. Teachers chosen to participate in the study were certified and received professional development about the KEEP model.

Instrumentation

Lunenburg and Irby (2008) affirmed, "Instrumentation is critical to adequately operationalize the variables of a research study" (p. 230). The instruments utilized in this study were surveys, which were created specifically for use in this study. The two surveys were used to collect data regarding teachers' and administrators' understanding

of the differences among the teacher performance ratings and of the calculation of the final summative teacher evaluation rating and to determine teachers' and administrators' perceptions of whether the district is using the KEEP model. Additionally, the survey was conducted to determine whether school level and whether a teacher had been evaluated affected teacher understanding and perceptions and whether school level affected administrator understanding and perceptions of the evaluation regarding KEEP.

Both surveys contained 34 items, separated into three sections. In the first section, all participants were asked to rate their level of understanding of the differences between the ratings on each of the teaching behaviors included on the KEEP tool. Participants were provided Likert-type scale responses of *do not understand, somewhat understand, mostly understand, and completely understand* the evaluation ratings related to the constructs of learner and learning, content knowledge, instructional practice, and professional responsibility (see Appendix A and B). In the second section of both surveys, participants were provided the selection choices of *strongly disagree, disagree, agree, strongly agree* that the district is using the KEEP model for continual improvement of instruction, meaningfully differentiating instruction, identify teacher needs, guiding professional development, and making informed personnel decisions. The last item in the section of both surveys teacher and administrator understanding of the calculation of the summative rating and if their understanding is affected by their school level (elementary [K-4], middle school [5-8], and high school [9-12]). The final section of the teacher survey was composed of two demographic items: school level and whether the teacher has been evaluated using KEEP. The final section of the

administrator survey was composed of one demographic item, school level (elementary [K-4], middle school [5-8], and high school [9-12], central office administrator).

Measurement. Items 1-27 in the teachers and administrators surveys were used to address the variables in RQ1-RQ7 regarding participant's level of understanding of the differences between the ratings for each of the teaching behaviors included on the KEEP tool. The format for items 1-27 was a 4-point Likert-type scale with the response options of *Do not Understand, Somewhat Understand, Mostly Understand, and Completely Understand* (see Appendices A and B). Table 1 contains the constructs included in the KEEP protocol with the associated survey items.

Table 1

Measurements for RQ1-RQ7

Construct	Items
Learner and Learning	
Learner development	Items 1, 2, and 3
Learner differences	Items 4 and 5
Learning environment	Items 6 and 7
Content Knowledge	
Content knowledge	Items 8 and 9
Innovative Application of content knowledge	Items 10, 11, and 12
Instructional Practice	
Planning for instruction	Items 13, 14, and 15
Assessment	Items 16, 17, and 18
Instructional strategies	Items 19, 20, and 21
Professional Responsibility	
Reflection and continuous growth	Items 22, 23, and 24
Collaboration and leadership	Items 25, 26, and 27

Survey items 28-32 were associated with the Kansas Educator Evaluation Protocol (KEEP) and the district use of the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions. The format for items 28-32 in both the teacher and administrator survey was a 4-point

Liker-type scale with the response options of S*trongly Disagree, Disagree, Agree, and Strongly Agree*. These items were used to address RQ15-RQ21 (see Appendix A).

The response options were *Elementary Teacher (K-4), Middle School Teacher (5-8), High School Teacher (9-12),* Teacher at multiple school levels, and other. In the teacher survey, the closed-ended item 33 was used to address RQ2, RQ7, RQ9, RQ14, RQ16, RQ19, and RQ21 (see Appendix A). Item 33 included the response options of *Central Office Administrator, Elementary Administrator (K-4), Middle School Administrator (5-8), and High School Administrator (9-12).* When teachers' and administrators' responses were compared, only three levels were included for administrators. In the administrator survey, the closed-ended item 33 was used to address RQ5, RQ7, RQ19, and RQ21 (see Appendix A).

In the teacher survey, the closed-ended item 34 was used to address RQ3, RQ10, and RQ17 to obtain the number of semesters teachers had been evaluated using KEEP. Item 34 within the teacher survey included possible response options of *I have not been evaluated using KEEP, Fall 2014, Spring 2015 and Fall 2015* (see Appendix *A*). *"I have not been evaluated using KEEP"* was recoded "Not evaluated." The other three responses were recoded "Evaluated."

Validity and reliability. To ensure the validity of the survey, input was sought from a panel of experts composed of teachers and administrators who did not work in District G. Each expert was provided a copy of the survey for review and recommendations. Modifications to the surveys were made according to recommendations from the expert panel.

A reliability analysis was not needed because a scale was not constructed from the survey items. The researcher used single-item measurement.

> Most commonly used single-item measures can be divided into two categories: (a) those measuring self-reported facts... and (b) those measuring psychological constructs, e.g., aspects of personality... measuring the former with single items is common practice. However, using a single-item measure for the latter is considered to be a "fatal error" in research. If the construct being measured is sufficiently narrow or is unambiguous to the respondent, a single item may suffice. (Sackett & Larson, 1990, p. 631)

The individual items used in this research were self-reported facts that were sufficiently narrow and unambiguous.

Data Collection Procedures

Prior to collecting the data, the researcher obtained permission to conduct the study from the Institutional Review Board of Baker University. An application, which included the survey instrument, was submitted to the Institutional Review Board of Baker University (see Appendix C). After permission was granted (see Appendix D), the District G directory was obtained online, which allowed access to District G's teacher and administrator email addresses. The directory was accessed online, and an email was sent to all teachers and administrators. Each teacher and administrator received a link to the online survey that included demographic information and knowledge based statements. A different link was created for administrators and teachers. The surveys were created using Google Forms. The researcher then utilized Google Forms. Included in the email used for the distribution of the survey were an informed consent letter and

the link to the survey (see Appendix E and F). At the closing of the survey, survey data were downloaded to an Excel spreadsheet. To complete the process, data was uploaded to IBM® SPSS® Statistics Faculty Pack 23 for Windows for data analysis.

Data Analysis and Hypothesis Testing

Twenty-one research questions guided this study. Each research question is paired with an associated hypothesis and statistical analysis.

RQ1. To what extent do teachers understand the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective?

Twenty-seven one-sample *t* tests were conducted to address RQ1. H1 through H27 were tested. For each test, the sample mean was compared to the reference value of 2.0 and the level of significance was set at .05**.**

H1. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning and alignment of instruction to meet student learning needs and developmental levels.

H2. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of teaching approaches and resources.

H3. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to adapting instruction to meet student needs.

H4. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to taking steps to gain knowledge of all students' individual differences.

H5. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using the knowledge of students to create a culture of respect among all students.

H6. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with students to promote student ownership of the learning.

H7. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to establishing a safe, respectful, and academically challenging environment.

H8. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to displaying knowledge of content by encouraging the use of multiple representations, explanations, and a wide variety of experiences.

H9. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using strategies to build student understanding of content.

H10. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using problem solving, critical thinking skills, and technology to explore and deliver content.

H11. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to exploring and delivering content through real world applications of knowledge.

H12. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with colleagues to provide cross curricular learning opportunities for students.

H13. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning rigorous and meaningful activities to meet the needs of all students.

H14. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using objectives that align with district, state, and/or national standards.

H15. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to meeting the needs of all students.

H16. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing opportunities for students to demonstrate learning through a variety of assessment methods.

H17. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using assessment data to inform instruction.

H18. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing assessment feedback to promote students' responsibility for their own learning.

H19. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of strategies to engage and challenge students in a variety of learning situations.

H20. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to incorporating strategies for differentiation and scaffolding for all students.

H21. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging students in higher order thinking.

H22. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging in ongoing and purposeful professional learning connected to student learning.

H23. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to reflecting on practice and actively seeks opportunities for improvement.

H24. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to analyzing and reflecting on student data to impact student growth.

H25. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with multiple stakeholders in school and professional activities.

H26. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of methods of communication with stakeholders.

H27. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to demonstrating leadership skills used to support and improve student learning.

RQ2. To what extent is teachers' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

Twenty-seven one-factor ANOVAs were conducted to address RQ2. H28 through H54 were tested. The categorical variable used to group the participants' responses to the survey was school level (elementary [K-4], middle school [5-8], and high school [9-12]). The level of significance was set at .05.

H28. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning and alignment of instruction to meet student learning needs and developmental levels is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H29. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of teaching approaches and resources is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H30. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to adapting instruction to

meet student needs is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H31. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to taking steps to gain knowledge of all students' individual differences is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H32. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using the knowledge of students to create a culture of respect among all students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H33. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with students to promote student ownership of the learning is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H34. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to establishing a safe, respectful, and academically challenging environment is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H35. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to displaying knowledge of content by encouraging the use of multiple representations, explanations, and a wide variety of experiences is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H36. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using strategies to build student understanding of content is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H37. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using problem solving, critical thinking skills, and technology to explore and deliver content is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H38. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to exploring and delivering content through real world applications of knowledge is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H39. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with colleagues to provide cross curricular learning opportunities for students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H40. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning rigorous and meaningful activities to meet the needs of all students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H41. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using objectives that

align with district, state, and/or national standards is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H42. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to meeting the needs of all students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H43. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing opportunities for students to demonstrate learning through a variety of assessment methods is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H44. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using assessment data to inform instruction is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H45. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing assessment feedback to promote students' responsibility for their own learning is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H46. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of strategies to engage and challenge students in a variety of learning situations is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H47. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to incorporating strategies for differentiation and scaffolding for all students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H48. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging students in higher order thinking is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H49. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging in ongoing and purposeful professional learning connected to student learning is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H50. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to reflecting on practice and actively seeks opportunities for improvement is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H51. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to analyzing and reflecting on student data to impact student growth is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H52. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with

multiple stakeholders in school and professional activities is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H53. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of methods of communication with stakeholders is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H54. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to demonstrating leadership skills used to support and improve student learning is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

RQ3. To what extent is teachers' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective affected by whether the teacher has been evaluated using KEEP?

Twenty-seven two-sample *t* tests were conducted to address RQ3. H55 through H81 were tested. For each test, the sample mean for teachers who were evaluated was compared with the sample mean for teachers who were not evaluated. The level of significance was set at .05.

H55. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning and alignment of instruction to meet student learning needs and developmental levels is affected by whether the teacher has been evaluated using KEEP.

H56. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of

teaching approaches and resources is affected by whether the teacher has been evaluated using KEEP.

H57. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to adapting instruction to meet student needs is affected by whether the teacher has been evaluated using KEEP.

H58. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to taking steps to gain knowledge of all students' individual differences is affected by whether the teacher has been evaluated using KEEP.

H59. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using the knowledge of students to create a culture of respect among all students is affected by whether the teacher has been evaluated using KEEP.

H60. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with students to promote student ownership of the learning is affected by whether the teacher has been evaluated using KEEP.

H61. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to establishing a safe, respectful, and academically challenging environment is affected by whether the teacher has been evaluated using KEEP.

H62. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to displaying knowledge

of content by encouraging the use of multiple representations, explanations, and a wide variety of experiences is affected by whether the teacher has been evaluated using KEEP.

H63. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using strategies to build student understanding of content is affected by whether the teacher has been evaluated using KEEP.

H64. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using problem solving, critical thinking skills, and technology to explore and deliver content is affected by whether the teacher has been evaluated using KEEP.

H65. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to exploring and delivering content through real world applications of knowledge is affected by whether the teacher has been evaluated using KEEP.

H66. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with colleagues to provide cross curricular learning opportunities for students is affected by whether the teacher has been evaluated using KEEP.

H67. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning rigorous and meaningful activities to meet the needs of all students is affected by whether the teacher has been evaluated using KEEP.

H68. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using objectives that align with district, state, and/or national standards is affected by whether the teacher has been evaluated using KEEP.

H69. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to meeting the needs of all students is affected by whether the teacher has been evaluated using KEEP.

H70. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing opportunities for students to demonstrate learning through a variety of assessment methods is affected by whether the teacher has been evaluated using KEEP.

H71. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using assessment data to inform instruction is affected by whether the teacher has been evaluated using KEEP.

H72. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing assessment feedback to promote students' responsibility for their own learning is affected by whether the teacher has been evaluated using KEEP.

H73. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of strategies to engage and challenge students in a variety of learning situations is affected by whether the teacher has been evaluated using KEEP.

H74. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to incorporating strategies for differentiation and scaffolding for all students is affected by whether the teacher has been evaluated using KEEP.

H75. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging students in high order thinking are affected by whether the teacher has been evaluated using KEEP.

H76. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging in ongoing and purposeful professional learning connected to student learning is affected by whether the teacher has been evaluated using KEEP.

H77. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to reflecting on practice and actively seeks opportunities for improvement is affected by whether the teacher has been evaluated using KEEP.

H78. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to analyzing and reflecting on student data to impact student growth is affected by whether the teacher has been evaluated using KEEP.

H79. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with multiple stakeholders in school and professional activities is affected by whether the teacher has been evaluated using KEEP.

H80. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of methods of communication with stakeholders is affected by whether the teacher has been evaluated using KEEP.

H81. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to demonstrating leadership skills used to support and improve student learning is affected by whether the teacher has been evaluated using KEEP.

RQ4. To what extent do administrators understand the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective?

Twenty-seven one-sample *t* tests were conducted to address RQ4. H82 through H108 were tested. For each test, the sample mean was compared to the reference value of 2.0 and the level of significance was set at .05**.**

H82. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning and alignment of instruction to meet student learning needs and developmental levels.

H83. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of teaching approaches and resources.

H84. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to adapting instruction to meet student needs.

H85. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to taking steps to gain knowledge of all students' individual differences.

H86. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using the knowledge of students to create a culture of respect among all students.

H87. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with students to promote student ownership of the learning.

H88. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to establishing a safe, respectful, and academically challenging environment.

H89. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to displaying knowledge of content by encouraging the use of multiple representations, explanations, and a wide variety of experiences.

H90. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using strategies to build student understanding of content.

H91. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using problem solving, critical thinking skills, and technology to explore and deliver content.

H92. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to exploring and delivering content through real world applications of knowledge.

H93. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with colleagues to provide cross curricular learning opportunities for students.

H94. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning rigorous and meaningful activities to meet the needs of all students.

H95. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using objectives that align with district, state, and/or national standards.

H96. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to meeting the needs of all students.

H97. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing opportunities for students to demonstrate learning through a variety of assessment methods.

H98. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using assessment data to inform instruction.

H99. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing assessment feedback to promote students' responsibility for their own learning.

H100. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of strategies to engage and challenge students in a variety of learning situations.

H101. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to incorporating strategies for differentiation and scaffolding for all students.

H102. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging students in higher order thinking.

H103. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging in ongoing and purposeful professional learning connected to student learning.

H104. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to reflecting on practice and actively seeks opportunities for improvement.

H105. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to analyzing and reflecting on student data to impact student growth.

H106. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with multiple stakeholders in school and professional activities.

H107. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of methods of communication with stakeholders.

H108. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to demonstrating leadership skills used to support and improve student learning.

RQ5. To what extent is administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12])?

Twenty-seven one-factor ANOVAs were conducted to address RQ5. H109 through H135 were tested. The categorical variable used to group the participants' responses to the survey was school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]). The level of significance was set at .05.

H109. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning and alignment of instruction to meet student learning needs and developmental levels is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H110. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of teaching approaches and resources is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H111. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to adapting instruction to meet student needs is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H112. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to taking steps to gain knowledge of all students' individual differences is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H113. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using the knowledge of students to create a culture of respect among all students is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H114. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with students to promote student ownership of the learning is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H115. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to establishing a safe, respectful, and academically challenging environment is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H116. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to displaying knowledge of content by encouraging the use of multiple representations, explanations, and a wide variety of experiences is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H117. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using strategies to build student understanding of content is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H118. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using problem solving, critical thinking skills, and technology to explore and deliver content is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H119. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to exploring and delivering content through real world applications of knowledge is affected by school

level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H120. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with colleagues to provide cross curricular learning opportunities for students is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H121. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning rigorous and meaningful activities to meet the needs of all students is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H122. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using objectives that align with district, state, and/or national standards is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H123. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to meeting the needs of all students is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H124. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing opportunities for students to demonstrate learning through a variety of assessment

methods is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H125. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using assessment data to inform instruction is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H126. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing assessment feedback to promote students' responsibility for their own learning is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H127. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of strategies to engage and challenge students in a variety of learning situations is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H128. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to incorporating strategies for differentiation and scaffolding for all students is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H129. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging

students in higher order thinking is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H130. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging in ongoing and purposeful professional learning connected to student learning is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H131. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to reflecting on practice and actively seeks opportunities for improvement is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H132. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to analyzing and reflecting on student data to impact student growth is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H133. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with multiple stakeholders in school and professional activities is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H134. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety

of methods of communication with stakeholders is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H135. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to demonstrating leadership skills used to support and improve student learning is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

RQ6. To what extent is there a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective?

Twenty-seven two-sample *t* tests were conducted to address RQ6. H136 through H162 were tested. The sample mean of the teachers' responses was compared to the sample mean of the administrators' responses. The level of significance was set at .05.

H136. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to planning and alignment of instruction to meet student learning needs and developmental levels.

H137. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using a variety of teaching approaches and resources.

H138. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to adapting instruction to meet student needs.

H139. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to taking steps to gain knowledge of all students' individual differences.

H140. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using the knowledge of students to create a culture of respect among all students.

H141. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to collaborating with students to promote student ownership of the learning.

H142. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to establishing a safe, respectful, and academically challenging environment.

H143. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to displaying knowledge of content by encouraging the use of multiple representations, explanations, and a wide variety of experiences.

H144. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective,

developing, or ineffective related to using strategies to build student understanding of content.

H145. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using problem solving, critical thinking skills, and technology to explore and deliver content.

H146. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to exploring and delivering content through real world applications of knowledge.

H147. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to collaborating with colleagues to provide cross curricular learning opportunities for students.

H148. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to planning rigorous and meaningful activities to meet the needs of all students.

H149. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using objectives that align with district, state, and/or national standards.

H150. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to meeting the needs of all students.

H151. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to providing opportunities for students to demonstrate learning through a variety of assessment methods.

H152. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using assessment data to inform instruction.

H153. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to providing assessment feedback to promote students' responsibility for their own learning.

H154. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using a variety of strategies to engage and challenge students in a variety of learning situations.

H155. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to incorporating strategies for differentiation and scaffolding for all students.

H156. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to engaging students in higher order thinking.

H157. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to engaging in ongoing and purposeful professional learning connected to student learning.

H158. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to reflecting on practice and actively seeks opportunities for improvement.

H159. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to analyzing and reflecting on student data to impact student growth.

H160. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to collaborating with multiple stakeholders in school and professional activities.

H161. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using a variety of methods of communication with stakeholders.

H162. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to demonstrating leadership skills used to support and improve student learning.

RQ7. To what extent is the difference between teachers' and administrators' understanding of the differences between highly effective, effective, developing, and ineffective ratings of teacher performance affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

Twenty-seven two-factor ANOVAs were conducted to address RQ7. H163 through H189 were tested. The two categorical variables were participant position (teacher and administrator) and school level (elementary [K-4], middle school [5-8], and high school [9-12]). The interaction between participant position and school level was used to test each hypothesis. The level of significance was set at .05.

H163. The difference between teachers' and administrators' understanding of the differences between highly effective, effective, developing, and ineffective ratings of teacher performance related to planning and alignment of instruction to meet student learning needs and developmental levels is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H164. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using a variety of teaching approaches and resources is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H165. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to adapting instruction to meet student needs is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H166. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to taking steps to gain knowledge of all students' individual differences is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H167. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using the knowledge of students to create a culture of respect among all students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H168. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to collaborating with students to promote student ownership of the learning is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H169. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to establishing a safe, respectful, and academically

challenging environment is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H170. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to displaying knowledge of content by encouraging the use of multiple representations, explanations, and a wide variety of experiences is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H171. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using strategies to build student understanding of content is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H172. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using problem solving, critical thinking skills, and technology to explore and deliver content is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H173. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to exploring and delivering content through real world applications of knowledge is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H174. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to collaborating with colleagues to provide cross curricular learning opportunities for students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H175. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to planning rigorous and meaningful activities to meet the needs of all students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H176. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using objectives that align with district, state, and/or national standards is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H177. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to meeting the needs of all students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H178. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to providing opportunities for students to demonstrate

learning through a variety of assessment methods is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H179. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using assessment data to inform instruction is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H180. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to providing assessment feedback to promote students' responsibility for their own learning is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H181. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using a variety of strategies to engage and challenge students in a variety of learning situations is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H182. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to incorporating strategies for differentiation and scaffolding for all students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H183. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective,

developing, or ineffective related to engaging students in higher order thinking is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H184. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to engaging in ongoing and purposeful professional learning connected to student learning is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H185. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to reflecting on practice and actively seeks opportunities for improvement is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H186. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to analyzing and reflecting on student data to impact student growth is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H187. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to collaborating with multiple stakeholders in school and professional activities is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H188. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using a variety of methods of communication with stakeholders is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H189. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to demonstrating leadership skills used to support and improve student learning is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

RQ8. To what extent do teachers understand the calculation of the final summative teacher evaluation rating?

H190. Teachers perceive that they understand the calculation of the final summative teacher evaluation rating.

A one-sample *t* test was conducted to address RQ8. H190 was tested. The sample mean was compared to the reference value of 2.0 and the level of significance was set at .05.

RQ9. To what extent is teachers' understanding of the calculation of the final summative teacher evaluation rating affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

H191. Teachers' understanding of the final summative teacher evaluation rating is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

A one-factor ANOVA was conducted to address RQ9. H191 was tested. The categorical variable used to group the participants' responses to the survey was school level (elementary [K-4], middle school [5-8], and high school [9-12]). The level of significance was set at .05.

RQ10. To what extent is teachers' understanding of the calculation of the final summative teacher evaluation rating affected by whether the teacher has been evaluated using KEEP?

H192. Teachers' understanding of the calculation of the final summative teacher evaluation rating is affected by whether the teacher has been evaluated using KEEP.

A two-sample *t* test was conducted to address RQ10. The sample mean for teachers who had been evaluated was compared to the sample mean for teachers who had not been evaluated. The level of significance was set at .05.

RQ11. To what extent do administrators understand the calculation of the final summative teacher evaluation rating?

H193. Administrators understand the calculation of the final summative teacher evaluation rating.

A one-sample *t* test was conducted to address RQ11. H193 was tested. The sample mean was compared to the reference value of 2.0 and the level of significance was set at .05.

RQ12. To what extent is administrators' understanding of the calculation of the final summative teacher evaluation rating affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

H194. Administrators' understanding of the final summative teacher evaluation rating is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

A one-factor ANOVA was conducted to address RQ12. H194 was tested. The categorical variable used to group the participants' responses to the survey was school level (elementary [K-4], middle school [5-8], and high school [9-12]). The level of significance was set at .05.

RQ13. To what extent is there a difference between teachers' and administrators' understanding of the calculation of the final summative teacher evaluation rating?

H195. There is a difference between teachers' and administrators' understanding of the calculation of the final summative teacher evaluation rating.

A two-sample *t* test was conducted to address RQ13. H195 was tested. The sample mean of the teachers' responses was compared to the sample mean of the administrators' responses. The level of significance was set at .05.

RQ14. To what extent is the difference between teachers' and administrators' understanding of the calculation of the final summative teacher evaluation rating affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

H196. The difference between teachers' and administrators' understanding of the calculation of the final summative teacher evaluation rating is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

A two-factor ANOVA was conducted to address RQ14. H196 was tested. The two categorical variables were participant position (teacher and administrator) and school level (elementary [K-4], middle school [5-8], and high school [9-12]). The interaction

between participant position and school level was used to test each hypothesis. The level of significance was set at .05.

RQ15. To what extent do teachers perceive that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions?

H197. Teachers perceive that the district is using the KEEP Model for continual improvement of instruction.

H198. Teachers perceive that the district is using the KEEP Model to meaningfully differentiate performance.

H199. Teachers perceive that the district is using the KEEP Model to provide useful feedback that identifies needs and guides professional development.

H200. Teachers perceive that the district is using the KEEP Model to make informed personnel decisions.

Four one-sample *t* tests were conducted to address RQ15. H197 through H200 were tested. For each test, the sample mean was compared to the reference value of 2.0 and the level of significance was set at .05.

RQ16. To what extent are teachers' perceptions that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

H201. Teachers' perceptions that the district is using the KEEP Model for continual improvement of instruction are affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H202. Teachers' perceptions that the district is using the KEEP Model to meaningfully differentiate performance are affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H203. Teachers' perceptions that the district is using the KEEP Model to provide useful feedback that identifies needs and guides professional development are affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]). .

H204. Teachers' perceptions that the district is using the KEEP Model to make informed personnel decisions are affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

Four one-factor ANOVAs were conducted to address RQ16. H201 through H204 were tested. The categorical variable used to group the participants' responses to the survey was school level (elementary [K-4], middle school [5-8], and high school [9-12]). The level of significance was set at .05.

RQ17. To what extent are teachers' perceptions that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions affected by whether the teacher has been evaluated using KEEP?

H205. Teachers' perceptions that the district is using the KEEP Model for continual improvement of instruction are affected by whether the teacher has been evaluated using KEEP.

H206. Teachers' perceptions that the district is using the KEEP Model to meaningfully differentiate performance are affected by whether the teacher has been evaluated using KEEP.

H207. Teachers' perceptions that the district is using the KEEP Model to provide useful feedback that identifies needs and guides professional development are affected by whether the teacher has been evaluated using KEEP.

H208. Teachers' perceptions that the district is using the KEEP Model to make informed personnel decisions are affected by whether the teacher has been evaluated using KEEP.

Four two-sample *t* tests were conducted to address RQ17. H205 through H208 were tested. For each test, the sample mean for teachers who were evaluated was compared with the sample mean for teachers who were not evaluated. The level of significance was set at .05.

RQ18. To what extent do administrators perceive that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions?

H209. Administrators perceive that the district is using the KEEP Model for continual improvement of instruction.

H210. Administrators perceive that the district is using the KEEP Model to meaningfully differentiate performance.

H211. Administrators perceive that the district is using the KEEP Model to provide useful feedback that identifies needs and guides professional development.

H212. Administrators perceive that the district is using the KEEP Model to make informed personnel decisions.

Four one-sample *t* tests were conducted to address RQ18. H209 through H212 were tested. For each test, the sample mean was compared to the reference value of 2.0 and the level of significance was set at .05.

RQ19. To what extent are administrators' perceptions that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12])?

H213. Administrators' perceptions that the district is using the KEEP Model for continual improvement of instruction are affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H214. Administrators' perceptions that the district is using the KEEP Model to meaningfully differentiate performance are affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H215. Administrators' perceptions that the district is using the KEEP Model to provide useful feedback that identifies needs and guides professional development are

affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H216. Administrators' perceptions that the district is using the KEEP Model to make informed personnel decisions are affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

Four one-factor ANOVAs were conducted to address RQ19. H213 through H216 were tested. The categorical variable used to group the participants' responses to the survey was school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]). The level of significance was set at .05.

RQ20. To what extent is there a difference between teachers' and administrators' perceptions that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions?

H217. There is a difference between teachers' and administrators' perceptions that the district is using the KEEP Model for continual improvement of instruction.

H218. There is a difference between teachers' and administrators' perceptions that the district is using the KEEP Model to meaningfully differentiate performance.

H219. There is a difference between teachers' and administrators' perceptions that the district is using the KEEP Model to provide useful feedback that identifies needs and guides professional development.

H220. There is a difference between teachers' and administrators' perceptions that the district is using the KEEP Model to make informed personnel decisions.

Four two-sample *t* tests were conducted to address RQ20. H217 through H220 were tested. The sample mean of the teachers' responses was compared to the sample mean of the administrators' responses. The level of significance was set at .05.

RQ21. To what extent is the difference between teachers' and administrators' perceptions that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

H221. The difference between teachers' and administrators' perceptions that the district is using the KEEP Model for continual improvement of instruction is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H222. The difference between teachers' and administrators' perceptions that the district is using the KEEP Model to meaningfully differentiate performance is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H223. The difference between teachers' and administrators' perceptions that the district is using the KEEP Model to provide useful feedback that identifies needs and guides professional development is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H224. The difference between teachers' and administrators' perceptions that the district is using the KEEP Model to make informed personnel decisions is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

Four two-factor ANOVAs were conducted to address RQ21. H221 through H224 were tested. The two categorical variables were participant position (teacher and

administrator) and school level (elementary [K-4], middle school [5-8], and high school [9-12]). The interaction between participant position and school level was used to test each hypothesis. The level of significance was set at .05.

Limitations

Lunenburg and Irby (2008) stated the limitations of a study are those factors that may have an effect on the explanation of the findings or generalizability of the results and are not under the control of the researcher. Potential limitations existing within this study included the response rate for individual survey items and the sample response rate. Limitations could also include the respondents' ability to comprehend and follow survey directions. The respondents experience with the KEEP could also be a limitation of the study.

Summary

Chapter three included the research design of this quantitative study. The participants were administrators and teachers within District G. The data collection and analysis procedures were discussed for each of the 21 research questions and 224 hypotheses. The limitations of the study were presented. Chapter four provides the results of the hypothesis testing and presents the calculation of the descriptive statistics.

Chapter Four

Results

Chapter four includes the results of the quantitative data analysis used to address the 21 research questions. The findings are presented beginning with an explanation of the descriptive statistics. Following the descriptive statistics, the results of the hypothesis tests are presented. The chapter ends with a summary.

Descriptive Statistics

A GoogleDoc link to the KEEP survey for teachers and administrators was sent to 438 certified teachers and 33 administrators within district G. Of those administrator and teacher invitations to complete the survey that were sent, 176 teachers and 14 administrators responded. The return rate of the survey was 40% for teachers and 45% for administrators. Of the surveys returned from teachers, 40% were from the elementary school levels, 24% were from middle school levels, and 36% were from the high school level. Of the surveys returned from administrators, 43% were from elementary school level, 14% were from middle school level, 14% were from the high school level, and 29% were from the central office.

Hypothesis Testing

The results of the hypothesis testing to address the 21 research questions presented in this study are discussed in this section. Each research question is followed by the method used to test each hypothesis. Next, the corresponding hypothesis statements are listed along with the results of each test. The significance level of .05 was utilized for all statistical analyses. The surveys can be found in Appendices A and B.

RQ1. To what extent do teachers understand the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective?

Twenty-seven one-sample *t* tests were conducted to address RQ1. H1 through H27 were tested. For each test, the sample mean was compared to the reference value of 2.0. The hypotheses that address RQ1 are listed below. The results of the hypothesis testing follow the hypotheses.

H1. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning and alignment of instruction to meet student learning needs and developmental levels.

H2. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of teaching approaches and resources.

H3. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to adapting instruction to meet student needs.

H4. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to taking steps to gain knowledge of all students' individual differences.

H5. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using the knowledge of students to create a culture of respect among all students.

H6. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with students to promote student ownership of the learning.

H7. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to establishing a safe, respectful, and academically challenging environment.

H8. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to displaying knowledge of content by encouraging the use of multiple representations, explanations, and a wide variety of experiences.

H9. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using strategies to build student understanding of content.

H10. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using problem solving, critical thinking skills, and technology to explore and deliver content.

H11. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to exploring and delivering content through real world applications of knowledge.

H12. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with colleagues to provide cross curricular learning opportunities for students.

H13. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning rigorous and meaningful activities to meet the needs of all students.

H14. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using objectives that align with district, state, and/or national standards.

H15. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to meeting the needs of all students.

H16. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing opportunities for students to demonstrate learning through a variety of assessment methods.

H17. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using assessment data to inform instruction.

H18. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing assessment feedback to promote students' responsibility for their own learning.

H19. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of strategies to engage and challenge students in a variety of learning situations.

H20. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to incorporating strategies for differentiation and scaffolding for all students.

H21. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging students in higher order thinking.

H22. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging in ongoing and purposeful professional learning connected to student learning.

H23. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to reflecting on practice and actively seeks opportunities for improvement.

H24. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to analyzing and reflecting on student data to impact student growth.

H25. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with multiple stakeholders in school and professional activities.

H26. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of methods of communication with stakeholders.

H27. Teachers understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to demonstrating leadership skills used to support and improve student learning.

The results of the hypothesis tests indicated that for each hypothesis the mean was statistically different from 2.0. See Table 2 for the hypothesis test statistics and the descriptive statistics. The means ranged from a minimum of 2.966 to a maximum of 3.534. The means indicated that the teachers mostly understood or completely understood the differences among the performance ratings of highly effective, effective, developing, or ineffective related to the evaluation constructs.

Table 2

Hypothesis Test and Descriptive Statistics for RQ1

Hypotheses		*t*	*df*	*p*	*M*	*SD*
H1	Alignment of Instruction	22.143	175	.000	3.193	.715
H2	Variety of Teaching Approaches	26.504	175	.000	3.403	.702
H3	Adapting Instruction	27.208	173	.000	3.374	.666
H4	Knowledge of Differences	17.263	175	.000	3.085	.834
H5	Culture of Respect	21.361	175	.000	3.216	.755
H6	Collaborating with Students	20.323	175	.000	3.210	.790
H7	Establishing an Environment	30.140	175	.000	3.534	.675
H8	Multiple Representations	19.929	175	.000	3.148	.764
H9	Strategies for Understanding	21.409	175	.000	3.250	.775
H10	Problem Solving and Tech.	20.242	175	.000	3.199	.786
H11	Real World Applications	21.617	175	.000	3.222	.750
H12	Cross Curricular Learning	22.663	175	.000	3.296	.758
H13	Rigorous Activities	22.788	175	.000	3.284	.748
H14	Alignment with Standards	21.915	175	.000	3.375	.832
H15	Meeting Student Needs	23.525	175	.000	3.335	.753
H16	Assessment Methods	24.270	175	.000	3.335	.730
H17	Assessment Data	24.095	175	.000	3.318	.726
H18	Assessment Feedback	19.798	175	.000	3.188	.796
H19	Strategies to Engage/Challenge	22.319	175	.000	3.256	.746
H20	Differentiation and Scaffolding	18.105	175	.000	3.102	.808
H21	Higher Order Thinking	20.756	175	.000	3.188	.759
H22	Professional Learning	22.139	175	.000	3.261	.756
H23	Seeks Opportunities	24.658	175	.000	3.347	.725
H24	Impact of Data on Learning	19.877	175	.000	3.199	.800
H25	Collaborate with Stakeholders	14.236	175	.000	2.966	.900
H26	Communication Methods	14.337	175	.000	2.966	.894
H27	Leadership Skills	19.685	175	.000	3.171	.789

RQ2. To what extent is teachers' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

Twenty-seven one-factor ANOVAs were conducted to address RQ2. H28 through H54 were tested. The categorical variable used to group the participants' responses to the survey was school level (elementary [K-4], middle school [5-8], and high school [9-12]). The results of the hypothesis testing follow the hypotheses.

H28. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning and alignment of instruction to meet student learning needs and developmental levels is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H29. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of teaching approaches and resources is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H30. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to adapting instruction to meet student needs is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H31. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to taking steps to gain knowledge of all students' individual differences is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H32. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using the knowledge of students to create a culture of respect among all students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H33. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with students to promote student ownership of the learning is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H34. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to establishing a safe, respectful, and academically challenging environment is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H35. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to displaying knowledge of content by encouraging the use of multiple representations, explanations, and a wide variety of experiences is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H36. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using strategies to build student understanding of content is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H37. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using problem solving,

critical thinking skills, and technology to explore and deliver content is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H38. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to exploring and delivering content through real world applications of knowledge is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H39. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with colleagues to provide cross curricular learning opportunities for students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H40. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning rigorous and meaningful activities to meet the needs of all students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H41. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using objectives that align with district, state, and/or national standards is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H42. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to meeting the needs of all students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H43. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing opportunities for students to demonstrate learning through a variety of assessment methods is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H44. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using assessment data to inform instruction is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H45. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing assessment feedback to promote students' responsibility for their own learning is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H46. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of strategies to engage and challenge students in a variety of learning situations is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H47. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to incorporating strategies for differentiation and scaffolding for all students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H48. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging students in

higher order thinking is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H49. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging in ongoing and purposeful professional learning connected to student learning is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H50. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to reflecting on practice and actively seeks opportunities for improvement is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H51. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to analyzing and reflecting on student data to impact student growth is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H52. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with multiple stakeholders in school and professional activities is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H53. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of methods of communication with stakeholders is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H54. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to demonstrating leadership skills used to support and improve student learning is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

The results of the hypothesis tests indicated that for each hypothesis none of the means was statistically different from the other means. See Table G1 in Appendix G for the hypothesis test statistics for RQ2. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective was not affected by school level.

RQ3. To what extent is teachers' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective affected by whether the teacher has been evaluated using KEEP?

Twenty-seven two-sample *t* tests were conducted to address RQ3. H55 through H81 were tested. For each test, the sample mean for teachers who were evaluated was compared with the sample mean for teachers who were not evaluated. The hypotheses that address research question three are listed below. The results of the hypothesis testing follow the hypotheses.

H55. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning and alignment of instruction to meet student learning needs and developmental levels is affected by whether the teacher has been evaluated using KEEP.

H56. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of

teaching approaches and resources is affected by whether the teacher has been evaluated using KEEP.

H57. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to adapting instruction to meet student needs is affected by whether the teacher has been evaluated using KEEP.

H58. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to taking steps to gain knowledge of all students' individual differences is affected by whether the teacher has been evaluated using KEEP.

H59. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using the knowledge of students to create a culture of respect among all students is affected by whether the teacher has been evaluated using KEEP.

H60. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with students to promote student ownership of the learning is affected by whether the teacher has been evaluated using KEEP.

H61. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to establishing a safe, respectful, and academically challenging environment is affected by whether the teacher has been evaluated using KEEP.

H62. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to displaying knowledge

of content by encouraging the use of multiple representations, explanations, and a wide variety of experiences is affected by whether the teacher has been evaluated using KEEP.

H63. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using strategies to build student understanding of content is affected by whether the teacher has been evaluated using KEEP.

H64. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using problem solving, critical thinking skills, and technology to explore and deliver content is affected by whether the teacher has been evaluated using KEEP.

H65. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to exploring and delivering content through real world applications of knowledge is affected by whether the teacher has been evaluated using KEEP.

H66. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with colleagues to provide cross curricular learning opportunities for students is affected by whether the teacher has been evaluated using KEEP.

H67. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning rigorous and meaningful activities to meet the needs of all students is affected by whether the teacher has been evaluated using KEEP.

H68. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using objectives that align with district, state, and/or national standards is affected by whether the teacher has been evaluated using KEEP.

H69. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to meeting the needs of all students is affected by whether the teacher has been evaluated using KEEP.

H70. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing opportunities for students to demonstrate learning through a variety of assessment methods is affected by whether the teacher has been evaluated using KEEP.

H71. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using assessment data to inform instruction is affected by whether the teacher has been evaluated using KEEP.

H72. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing assessment feedback to promote students' responsibility for their own learning is affected by whether the teacher has been evaluated using KEEP.

H73. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of strategies to engage and challenge students in a variety of learning situations is affected by whether the teacher has been evaluated using KEEP.

H74. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to incorporating strategies for differentiation and scaffolding for all students is affected by whether the teacher has been evaluated using KEEP.

H75. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging students in high order thinking are affected by whether the teacher has been evaluated using KEEP.

H76. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging in ongoing and purposeful professional learning connected to student learning is affected by whether the teacher has been evaluated using KEEP.

H77. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to reflecting on practice and actively seeks opportunities for improvement is affected by whether the teacher has been evaluated using KEEP.

H78. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to analyzing and reflecting on student data to impact student growth is affected by whether the teacher has been evaluated using KEEP.

H79. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with multiple stakeholders in school and professional activities is affected by whether the teacher has been evaluated using KEEP.

H80. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of methods of communication with stakeholders is affected by whether the teacher has been evaluated using KEEP.

H81. Teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to demonstrating leadership skills used to support and improve student learning is affected by whether the teacher has been evaluated using KEEP.

The results of three of the two-sample *t* tests were statistically significant. The results of the other 24 tests were not statistically significant. The statistical findings for all of the tests are found in Appendix G, Table G2. The test for item 20 indicated a statistically significant difference between the two means, $t = 2.697$, $df = 173$, $p = .008$. The sample mean for teachers who had been evaluated ($M = 3.187$, $SD = .758$) was higher than the sample mean for teachers who had not been evaluated ($M = 2.805$, $SD = .901$). On average, teachers who had been evaluated better understood the differences in the ratings of their performance on incorporating strategies for differentiation and scaffolding for all students than teachers who had not been evaluated. The test for item 22 indicated a statistically significant difference between the two means, $t = 2.035$, $df = 173$, $p = .043$. The sample mean for teachers who had been evaluated ($M = 3.321$, $SD = .700$) was higher than the sample mean for teachers who had not been evaluated ($M = 3.049$, $SD = .893$). On average, teachers who had been evaluated better understood differences in the ratings of their performance on engaging in ongoing and purposeful professional learning connected to student learning than teachers who had not been

evaluated. The test for item 23 indicated a statistically significant difference between the two means, $t = 2.774$, $df = 173$, $p = .006$. The sample mean for teachers who had been evaluated ($M = 3.425$, $SD = .665$) was higher than the sample mean for teachers who had not been evaluated ($M = 3.073$, $SD = .848$). On average, teachers who had been evaluated better understood differences in the ratings of their performance on reflecting on practice and actively seeks opportunities for improvement than teachers who had not been evaluated.

RQ4. To what extent do administrators understand the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective?

Twenty-seven one-sample *t* tests were conducted to address RQ4. H82 through H108 were tested. For each test, the sample mean was compared to the reference value of 2.0**.** The results of the hypothesis testing follow the hypotheses.

H82. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning and alignment of instruction to meet student learning needs and developmental levels.

H83. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of teaching approaches and resources.

H84. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to adapting instruction to meet student needs.

H85. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to taking steps to gain knowledge of all students' individual differences.

H86. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using the knowledge of students to create a culture of respect among all students.

H87. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with students to promote student ownership of the learning.

H88. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to establishing a safe, respectful, and academically challenging environment.

H89. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to displaying knowledge of content by encouraging the use of multiple representations, explanations, and a wide variety of experiences.

H90. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using strategies to build student understanding of content.

H91. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using problem solving, critical thinking skills, and technology to explore and deliver content.

H92. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to exploring and delivering content through real world applications of knowledge.

H93. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with colleagues to provide cross curricular learning opportunities for students.

H94. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning rigorous and meaningful activities to meet the needs of all students.

H95. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using objectives that align with district, state, and/or national standards.

H96. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to meeting the needs of all students.

H97. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing opportunities for students to demonstrate learning through a variety of assessment methods.

H98. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using assessment data to inform instruction.

H99. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing assessment feedback to promote students' responsibility for their own learning.

H100. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of strategies to engage and challenge students in a variety of learning situations.

H101. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to incorporating strategies for differentiation and scaffolding for all students.

H102. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging students in higher order thinking.

H103. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging in ongoing and purposeful professional learning connected to student learning.

H104. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to reflecting on practice and actively seeks opportunities for improvement.

H105. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to analyzing and reflecting on student data to impact student growth.

H106. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with multiple stakeholders in school and professional activities.

H107. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of methods of communication with stakeholders.

H108. Administrators understand the differences among the performance ratings of highly effective, effective, developing, or ineffective related to demonstrating leadership skills used to support and improve student learning.

The results of the hypothesis tests indicated that for each hypothesis the mean was statistically different from 2.0. See Table 3 for the hypothesis test statistics. The means ranged from a minimum of 3.071 to a maximum of 3.500. Each of the means indicated that the administrators mostly understood or completely understood the differences among the performance ratings of highly effective, effective, developing, or ineffective.

Table 3

Hypothesis Test and Descriptive Statistics for RQ4

Hypotheses		*t*	*df*	*p*	*M*	*SD*
H82	Alignment of Instruction	6.497	13	.000	3.214	.700
H83	Variety Teaching Approaches	8.018	13	.000	3.357	.633
H84	Adapting Instruction	8.629	13	.000	3.500	.650
H85	Knowledge of Differences	6.450	13	.000	3.143	.663
H86	Culture of Respect	5.551	13	.000	3.143	.770
H87	Collaborating with Students	6.511	13	.000	3.071	.616
H88	Establishing an Environment	8.629	13	.000	3.500	.650
H89	Multiple Representations	6.624	13	.000	3.286	.726
H90	Strategies for Understanding	8.272	13	.000	3.429	.646
H91	Problem Solving and Tech.	6.497	13	.000	3.214	.699
H92	Real World Applications	6.511	13	.000	3.071	.616
H93	Cross Curricular Learning	5.667	13	.000	3.214	.802
H94	Rigorous Activities	7.389	13	.000	3.500	.760
H95	Alignment with Standards	7.389	13	.000	3.500	.760
H96	Meeting Student Needs	8.272	13	.000	3.429	.646
H97	Assessment Methods	8.018	13	.000	3.357	.633
H98	Assessment Data	10.212	13	.000	3.357	.497
H99	Assessment Feedback	7.848	13	.000	3.214	.579
H100	Strategies to Engage/Challenge	6.624	13	.000	3.286	.726
H101	Differentiation and Scaffolding	5.551	13	.000	3.143	.770
H102	Higher Order Thinking	6.817	13	.000	3.357	.745
H103	Professional Learning	10.408	13	.000	3.429	.514
H104	Seeks Opportunities	6.817	13	.000	3.357	.745
H105	Impact of Data on Learning	6.624	13	.000	3.286	.726
H106	Collaborate with Stakeholders	5.491	13	.000	3.071	.730
H107	Communication Methods	6.817	13	.000	3.357	.745
H108	Leadership Skills	5.491	13	.000	3.071	.730

RQ5. To what extent is administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12])?

The twenty-seven one-factor ANOVAs to address RQ5 could not be conducted. H109 through H135 were not tested. Small sample sizes when the response data was disaggregated by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]) compromised the hypothesis testing.

H109. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning and alignment of instruction to meet student learning needs and developmental levels is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H110. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of teaching approaches and resources is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H111. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to adapting instruction to meet student needs is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H112. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to taking steps to

gain knowledge of all students' individual differences is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H113. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using the knowledge of students to create a culture of respect among all students is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H114. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with students to promote student ownership of the learning is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H115. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to establishing a safe, respectful, and academically challenging environment is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H116. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to displaying knowledge of content by encouraging the use of multiple representations, explanations, and a wide variety of experiences is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H117. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using strategies to build student understanding of content is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H118. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using problem solving, critical thinking skills, and technology to explore and deliver content is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H119. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to exploring and delivering content through real world applications of knowledge is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H120. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with colleagues to provide cross curricular learning opportunities for students is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H121. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to planning rigorous and meaningful activities to meet the needs of all students is affected by school

level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H122. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using objectives that align with district, state, and/or national standards is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H123. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to meeting the needs of all students is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H124. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing opportunities for students to demonstrate learning through a variety of assessment methods is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H125. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using assessment data to inform instruction is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H126. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to providing assessment feedback to promote students' responsibility for their own learning is affected

by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H127. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of strategies to engage and challenge students in a variety of learning situations is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H128. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to incorporating strategies for differentiation and scaffolding for all students is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H129. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging students in higher order thinking is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H130. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to engaging in ongoing and purposeful professional learning connected to student learning is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H131. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to reflecting on

practice and actively seeks opportunities for improvement is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H132. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to analyzing and reflecting on student data to impact student growth is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H133. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to collaborating with multiple stakeholders in school and professional activities is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H134. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to using a variety of methods of communication with stakeholders is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H135. Administrators' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective related to demonstrating leadership skills used to support and improve student learning is affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

RQ6. To what extent is there a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective?

Twenty-seven two-sample *t* tests were conducted to address RQ6. H136 through H162 were tested. The sample mean of the teachers' responses was compared to the sample mean of the administrators' responses. The results of the hypothesis testing follow the hypotheses.

H136. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to planning and alignment of instruction to meet student learning needs and developmental levels.

H137. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using a variety of teaching approaches and resources.

H138. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to adapting instruction to meet student needs.

H139. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to taking steps to gain knowledge of all students' individual differences.

H140. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective,

developing, or ineffective related to using the knowledge of students to create a culture of respect among all students.

H141. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to collaborating with students to promote student ownership of the learning.

H142. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to establishing a safe, respectful, and academically challenging environment.

H143. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to displaying knowledge of content by encouraging the use of multiple representations, explanations, and a wide variety of experiences.

H144. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using strategies to build student understanding of content.

H145. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using problem solving, critical thinking skills, and technology to explore and deliver content.

H146. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to exploring and delivering content through real world applications of knowledge.

H147. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to collaborating with colleagues to provide cross curricular learning opportunities for students.

H148. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to planning rigorous and meaningful activities to meet the needs of all students.

H149. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using objectives that align with district, state, and/or national standards.

H150. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to meeting the needs of all students.

H151. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to providing opportunities for students to demonstrate learning through a variety of assessment methods.

H152. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using assessment data to inform instruction.

H153. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to providing assessment feedback to promote students' responsibility for their own learning.

H154. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using a variety of strategies to engage and challenge students in a variety of learning situations.

H155. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to incorporating strategies for differentiation and scaffolding for all students.

H156. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to engaging students in higher order thinking.

H157. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to engaging in ongoing and purposeful professional learning connected to student learning.

H158. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to reflecting on practice and actively seeks opportunities for improvement.

H159. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to analyzing and reflecting on student data to impact student growth.

H160. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to collaborating with multiple stakeholders in school and professional activities.

H161. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using a variety of methods of communications with stakeholders.

H162. There is a difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to demonstrating leadership skills used to support and improve student learning.

The results of the hypothesis tests indicated that for each hypothesis the two means were not statistically different. See Table G3 in Appendix G for the hypothesis test statistics.

There were no statistically significant differences between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to the evaluation constructs.

RQ7. To what extent is the difference between teachers' and administrators' understanding of the differences between highly effective, effective, developing, and ineffective ratings of teacher performance affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

The 27 one-factor ANOVAs to address RQ7 could not be conducted. H163 through H189 were not tested. Small sample sizes when the response data was disaggregated by school level (elementary [K-4], middle school [5-8], and high school [9-12]) compromised the hypothesis testing.

H163. The difference between teachers' and administrators' understanding of the differences between highly effective, effective, developing, and ineffective ratings of teacher performance related to planning and alignment of instruction to meet student learning needs and developmental levels is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H164. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using a variety of teaching approaches and resources is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H165. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective,

developing, or ineffective related to adapting instruction to meet student needs is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H166. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to taking steps to gain knowledge of all students' individual differences is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H167. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using the knowledge of students to create a culture of respect among all students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H168. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to collaborating with students to promote student ownership of the learning is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H169. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to establishing a safe, respectful, and academically challenging environment is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H170. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to displaying knowledge of content by encouraging the use of multiple representations, explanations, and a wide variety of experiences is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H171. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using strategies to build student understanding of content is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H172. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using problem solving, critical thinking skills, and technology to explore and deliver content is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H173. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to exploring and delivering content through real world applications of knowledge is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H174. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to collaborating with colleagues to provide cross

curricular learning opportunities for students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H175. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to planning rigorous and meaningful activities to meet the needs of all students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H176. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using objectives that align with district, state, and/or national standards is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H177. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to meeting the needs of all students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H178. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to providing opportunities for students to demonstrate learning through a variety of assessment methods is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H179. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective,

developing, or ineffective related to using assessment data to inform instruction is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H180. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to providing assessment feedback to promote students' responsibility for their own learning is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H181. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using a variety of strategies to engage and challenge students in a variety of learning situations is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H182. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to incorporating strategies for differentiation and scaffolding for all students is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H183. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to engaging students in higher order thinking is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H184. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective,

developing, or ineffective related to engaging in ongoing and purposeful professional learning connected to student learning is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H185. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to reflecting on practice and actively seeks opportunities for improvement is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H186. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to analyzing and reflecting on student data to impact student growth is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H187. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to collaborating with multiple stakeholders in school and professional activities is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H188. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to using a variety of methods of communications with stakeholders is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H189. The difference between teachers' and administrators' understanding of the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective related to demonstrating leadership skills used to support and improve student learning is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

RQ8. To what extent do teachers understand the calculation of the final summative teacher evaluation rating?

H190. Teachers perceive that they understand the calculation of the final summative teacher evaluation rating.

A one-sample *t* test was conducted to address RQ8. H190 was tested. The sample mean was compared to the reference value of 2.0. The results of the one-sample *t* test indicated a statistically significant difference between the values, $t = 7.667$, $df = 175$, $p = .000$. The sample mean ($M = 2.506$, $SD = .875$) was higher than the reference value which indicated that the teachers mostly understood the calculation of the final summative teacher evaluation rating.

RQ9. To what extent is teachers' understanding of the calculation of the final summative teacher evaluation rating affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

H191. Teachers' understanding of the final summative teacher evaluation rating is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

A one-factor ANOVA was conducted to address RQ9. The categorical variable used to group teachers' understanding of the final summative teacher evaluation rating was school level (elementary [K-4], middle school [5-8], and high school [9-12]). The

results of the analyses indicated no statistically significant differences among the means, $F = 1.385$, $df = 2, 169$, $p = .253$. Teachers' understanding of the final summative teacher evaluation rating was not affected by school level.

RQ10. To what extent is teachers' understanding of the calculation of the final summative teacher evaluation rating affected by whether the teacher has been evaluated using KEEP?

H192. Teachers' understanding of the calculation of the final summative teacher evaluation rating is affected by whether the teacher has been evaluated using KEEP.

A two-sample *t* test was conducted to address RQ10. The two sample means were compared. The results of the two-sample *t* test indicated a statistically significant difference between the two values, $t = 2.581$, $df = 173$, $p = .011$. The sample mean for teachers who had been evaluated ($M = 2.590$, $SD = .869$) was higher than the sample mean for the teachers who had not been evaluated ($M = 2.195$, $SD = .813$). Teachers' understanding of the calculation of the final summative teacher evaluation rating was affected by whether the teacher had been evaluated using KEEP.

RQ11. To what extent do administrators understand the calculation of the final summative teacher evaluation rating?

H193. Administrators understand the calculation of the final summative teacher evaluation rating.

A two-sample *t* test was conducted to address RQ11. The two sample means were compared. The results of the one-sample *t* test indicated a statistically significant difference between the values, $t = 15$, $df = 13$, $p = .000$. The sample mean ($M = 3.071$, $SD = .267$) was higher than the reference value of 2.0, which indicated that the

administrators mostly understood the calculation of the final summative teacher evaluation rating.

RQ12. To what extent is administrators' understanding of the calculation of the final summative teacher evaluation rating affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

H194. Administrators' understanding of the final summative teacher evaluation rating is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

A one-factor ANOVA could not be conducted to address RQ12. H194 was not tested. Small sample sizes when the response data was disaggregated by school level (elementary [K-4], middle school [5-8], and high school [9-12]) compromised the hypothesis testing.

RQ13. To what extent is there a difference between teachers' and administrators' understanding of the calculation of the final summative teacher evaluation rating?

H195. There is a difference between teachers' and administrators' understanding of the calculation of the final summative teacher evaluation rating.

A two-sample *t* test was conducted to address RQ13. H195 was tested. The sample mean of the teachers' responses was compared to the sample mean of the administrators' responses. The results of the two-sample *t* test indicated a statistically significant difference between the two values, $t = 2.405$, $df = 188$, $p = .017$. The sample mean for the administrators was ($M = 3.071$, $SD = .267$) higher than the sample mean for the teachers ($M = 2.506$, $SD = .875$). On average, administrators understood the calculation of the final summative teacher evaluation better than the teachers did.

RQ14. To what extent is the difference between teachers' and administrators' understanding of the calculation of the final summative teacher evaluation rating affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

H196. The difference between teachers' and administrators' understanding of the calculation of the final summative teacher evaluation rating is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

The two-factor ANOVA to address RQ14 could not be conducted. H196 was not tested. Small sample sizes when the response data was disaggregated by school level (elementary [K-4], middle school [5-8], and high school [9-12]) compromised the hypothesis testing.

RQ15. To what extent do teachers perceive that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions?

Four one-sample *t* tests were conducted to address RQ15. H197 through H200 were tested. For each test, the sample mean was compared to the reference value of 2.0. The hypotheses that address RQ15 are listed below. The results of the hypothesis testing follow the hypotheses.

H197. Teachers perceive that the district is using the KEEP Model for continual improvement of instruction.

H198. Teachers perceive that the district is using the KEEP Model to meaningfully differentiate performance.

H199. Teachers perceive that the district is using the KEEP Model to provide useful feedback that identifies needs and guides professional development.

H200. Teachers perceive that the district is using the KEEP Model to make informed personnel decisions.

The results of the hypothesis tests indicated that for each hypothesis the mean was different from 2.0. See Table 4 for the hypothesis test statistics and descriptive statistics. The means ranged from a minimum of 2.716 to a maximum of 3.063. The means indicated that the teachers agreed that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions.

Table 4

Hypothesis Test and Descriptive Statistics for RQ15

Hypotheses	*t*	*df*	*p*	*M*	*SD*
H197 Continual Improvement of Instruction	23.622	175	.000	3.063	.597
H198 Meaningfully Differentiate Performance.	13.409	175	.000	2.716	.708
H199 Feedback for Professional Development	16.466	175	.000	2.864	.696
H200 Make Informed Personnel Decisions	17.163	175	.000	2.790	.610

RQ16. To what extent are teachers' perceptions that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

Four one-factor ANOVAs were conducted to address RQ16. H201- H204 were tested. The categorical variable used to group the participants' responses to the survey was school level (elementary [K-4], middle school [5-8], and high school [9-12]). The results of the hypothesis testing follow the hypotheses.

H201. Teachers' perceptions that the district is using the KEEP Model for continual improvement of instruction are affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H202. Teachers' perceptions that the district is using the KEEP Model to meaningfully differentiate performance are affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H203. Teachers' perceptions that the district is using the KEEP Model to provide useful feedback that identifies needs and guides professional development are affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]). .

H204. Teachers' perceptions that the district is using the KEEP Model to make informed personnel decisions are affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]). .

The results of the hypothesis tests for H201, H202, and H203 indicated that there was not a statistically significant difference between any of the means. See Table 5 for the hypothesis test statistics. For H204, there was a significant difference between at least two means, $F = 4.085$, $df = 2, 169$, $p = .019$. Teachers' perceptions that the district is using the KEEP Model to make informed personnel decisions were affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]). A follow-up post hoc test, the Tukey's HSD, indicated that the mean for elementary teachers ($M = 2.618$,

SD = .670) was significantly lower than the mean for middle school teachers (*M* =2.889, *SD* = .532). Elementary teachers did not agree with the middle school and high school teachers that the district is using the KEEP Model to make informed personnel decisions were affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

Table 5

Hypothesis Test Statistics for RQ16

Hypotheses	*df1*	*df2*	F	*p*
H201Continual Improvement	2	169	2.030	.134
H202 Differentiate Performance	2	169	2.225	.111
H203 Guide Professional Development	2	169	1.323	.269
H204 Make Informed Personnel Decisions	2	169	4.085	.019

RQ17. To what extent are teachers' perceptions that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions affected by whether the teacher has been evaluated using KEEP?

Four two-sample *t* tests were conducted to address RQ17. H205 through H208 were tested. For each test, the sample mean for teachers who were evaluated was compared with the sample mean for teachers who were not evaluated. The results of the hypothesis testing follow the hypotheses.

H205. Teachers' perceptions that the district is using the KEEP Model for continual improvement of instruction are affected by whether the teacher has been evaluated using KEEP.

H206. Teachers' perceptions that the district is using the KEEP Model to meaningfully differentiate performance are affected by whether the teacher has been evaluated using KEEP.

H207. Teachers' perceptions that the district is using the KEEP Model to provide useful feedback that identifies needs and guides professional development are affected by whether the teacher has been evaluated using KEEP.

H208. Teachers' perceptions that the district is using the KEEP Model to make informed personnel decisions are affected by whether the teacher has been evaluated using KEEP.

The results of the hypothesis tests indicated that for each hypothesis, the means were not statistically different based on whether the teacher was evaluated using KEEP. See Table 6 for the hypothesis test statistics. Teachers' perceptions that the district is using the KEEP Model to make informed personnel decisions, for continual improvement of instruction, to meaningfully differentiate performance, and provide useful feedback that identifies needs and guides professional development were not affected by whether the teacher has been evaluated using KEEP.

Table 6

Hypothesis Test Statistics for RQ17

Hypotheses	*t*	*df*	*p*
H205 Continual Improvement	.703	173	.483
H206 Differentiate Performance	.072	173	.943
H207 Guide Professional Development	1.446	173	.150
H208 Make Informed Personnel Decisions	.029	173	.977

RQ18. To what extent do administrators perceive that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions?

Four one-sample *t* tests were conducted to address RQ18. H209 through H212 were tested. For each test, the sample mean was compared to the reference value of 2.0. The results of the hypothesis testing follow the hypotheses.

H209. Administrators perceive that the district is using the KEEP Model for continual improvement of instruction?

H210. Administrators perceive that the district is using the KEEP Model to meaningfully differentiate performance.

H211. Administrators perceive that the district is using the KEEP Model to provide useful feedback that identifies needs and guides professional development.

H212. Administrators perceive that the district is using the KEEP Model to make informed personnel decisions.

The results of the hypothesis tests indicated that for each hypothesis the mean was statistically different from 2.0. See Table 7 for the hypothesis test statistics and descriptive statistics. The means ranged from a minimum of 2.571 to a maximum of 3.357. The means indicated that the administrators mostly agreed that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions.

Table 7

Hypothesis Test Statistics and Descriptive Statistics for RQ18

Hypotheses	*t*	*df*	*p*	*M*	*SD*
H209 Continual Improvement	8.018	13	.000	3.357	.633
H210 Differentiate Performance	4.163	13	.001	2.571	.514
H211 Guide Professional Development	4.372	13	.001	2.714	.311
H212 Make Informed Personnel Decisions	8.832	13	.000	2.857	.363

RQ19. To what extent are administrators' perceptions that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12])?

H213. Administrators' perceptions that the district is using the KEEP Model for continual improvement of instruction are affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H214. Administrators' perceptions that the district is using the KEEP Model to meaningfully differentiate performance are affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H215. Administrators' perceptions that the district is using the KEEP Model to provide useful feedback that identifies needs and guides professional development are affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

H216. Administrators' perceptions that the district is using the KEEP Model to make informed personnel decisions are affected by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]).

The one-factor ANOVAs to address RQ19 could not be conducted. H213-H216 were not tested. Small sample sizes when the response data was disaggregated by school level (central office administrator, elementary [K-4], middle school [5-8], and high school [9-12]) compromised the hypothesis testing.

RQ20. To what extent is there a difference between teachers' and administrators' perceptions that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions?

Four two-sample *t* tests were conducted to address RQ20. For each of hypothesis test, the two sample means were compared. The two sample means of the difference between teachers' and administrators' perceptions that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional

development, and to inform personnel decisions were compared. The results of the hypothesis testing follow the hypotheses.

H217. There is a difference between teachers' and administrators' perceptions that the district is using the KEEP Model for continual improvement of instruction.

H218. There is a difference between teachers' and administrators' perceptions that the district is using the KEEP Model to meaningfully differentiate performance.

H219. There is a difference between teachers' and administrators' perceptions that the district is using the KEEP Model to provide useful feedback that identifies needs and guides professional development.

H220. There is a difference between teachers' and administrators' perceptions that the district is using the KEEP Model to make informed personnel decisions.

The results of the two-sample *t* tests indicated no difference between teachers' and administrators' perceptions that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, or to inform personnel decisions. See Table 8 for the hypothesis test statistics.

Table 8

Hypothesis Test Statistics for RQ20

Hypotheses	*t*	*df*	*p*
H217 Make Informed Personnel Decisions	-1.770	188	.078
H218 Make Informed Personnel Decisions	.747	188	.456
H219 Make Informed Personnel Decisions	.779	188	.437
H220 Make Informed Personnel Decisions	-.407	188	.685

RQ21. To what extent is the difference between teachers' and administrators' perceptions that the district is using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions affected by school level (elementary [K-4], middle school [5-8], and high school [9-12])?

H221. The difference between teachers' and administrators' perceptions that the district is using the KEEP Model for continual improvement of instruction is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H222. The difference between teachers' and administrators' perceptions that the district is using the KEEP Model to meaningfully differentiate performance is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H223. The difference between teachers' and administrators' perceptions that the district is using the KEEP Model to provide useful feedback that identifies needs and guides professional development is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

H224. The difference between teachers' and administrators' perceptions that the district is using the KEEP Model to make informed personnel decisions is affected by school level (elementary [K-4], middle school [5-8], and high school [9-12]).

The two-factor ANOVAs to address RQ21 could not be conducted. H221- H224 were not tested. Small sample sizes when the response data was disaggregated by school level (elementary [K-4], middle school [5-8], and high school [9-12]) compromised the hypothesis testing.

Summary

Chapter four included the descriptive statistics and results of the hypothesis testing for this study. All 21 research questions and 224 hypotheses were reviewed, although six could not be conducted due to the sample size for administrators. Chapter five contains an overview of the study, an overview of the problem, the purpose statement and research questions, a review of the methodology, and the major findings. Implications for action, recommendations for future research, and concluding remarks are also presented.

Chapter Five

Interpretation and Recommendations

Teachers and administrators need to understand the differences among the teacher performance ratings and the calculation of the final summative rating. Understanding teacher evaluation is essential because it promotes professional development and allows for a clear assessment of teachers' and administrators' understandings regarding the evaluation process and professional development for improved instruction and student achievement. Chapter five includes a review of this study by providing an overview of the problem, purpose statement and research questions, review of methodology, and major findings. Additionally, the findings related to the literature, implications for action, recommendations for future research, and concluding remarks are contained within the chapter.

Study Summary

Teachers and administrators need to understand fully the evaluation tool used by a district because worthwhile professional development for teachers and administrators needs to be created to ensure that teachers are highly effective and students are learning. Understanding the evaluation tool may also lead to developing appropriate objectives for the mentoring programs for first and second-year teachers in a district. A common goal for district leadership, building administrators, and teachers is to develop highly effective educators. The evaluation model created by KSDE is KEEP, which is used by 72 districts and is the focus of this study. Provided in this section are an overview of the problem, purpose statement and research questions, review of the methodology, and major findings.

Overview of the problem. Without a clear assessment of teachers' and administrators' understandings of the evaluation process, creating a professional development plan that provides improved instruction and student achievement is difficult. At the time of this study, the newly implemented evaluation system had been in place for 1.5 years. The new evaluation process was implemented to provide better feedback to teachers to improve instruction and provide useful feedback that identifies needs and guides professional development. The new evaluation process also allowed for the integration of Student Growth Measures and evaluation goals.

Purpose statement and research questions. The first purpose of this study was to determine whether teachers and administrators understood the differences among the teacher performance ratings of highly effective, effective, developing, and ineffective. The second purpose of the study was to determine if teacher' understanding of the ratings were affected by school level and whether the teacher has been evaluated using KEEP and if administrators' understanding of the ratings were affected by school level. The third purpose of the study was to determine whether there were differences between teachers' and administrators' understanding of the ratings and whether the differences were affected by school level. The fourth purpose of the study was to determine teachers' and administrators' understanding of the calculation of the final summative teacher evaluation rating. The fifth purpose of the study was to determine if teacher' understanding of the calculation of the final summative teacher evaluation rating were affected by school level and whether the teacher has been evaluated using KEEP and if administrators' understanding of the final summative teacher evaluation rating were affected by school level. The sixth purpose of the study was to determine whether there

were differences between teachers' and administrators' understanding the final summative teacher evaluation rating and whether the differences were affected by school level. The seventh purpose of the study was to determine teachers' and administrators' perceptions of whether the district is using the KEEP model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and it inform personnel decisions. Twenty-one research questions were established to address the purposes of the study.

Review of the methodology. A descriptive non-experimental survey research design was conducted during the 2015-2016 school year. Teachers and administrators who used the Kansas Education Evaluation Protocol in a public school district in Johnson County, Kansas were the participants in this study. A survey created by the researcher was used to collect data for the study. Quantitative data analysis was used to test the 224 hypotheses that addressed the 21 research questions.

Major findings. Various major findings were identified in the current research study. All major findings are grouped by the twenty-one research questions. Each major finding is listed in the order of the research questions.

Twenty-seven hypotheses were tested to address RQ1. The results of the hypotheses testing indicated that teachers mostly or completely understood the differences among the performance ratings of highly effective, effective, developing, or ineffective. Additionally, the results of the hypothesis testing for 27 hypotheses related to RQ2 indicated that teachers' understanding of the differences among the performance ratings of highly effective, effective, developing, or ineffective were not affected by school level. The results of three of the 27 hypotheses tests related to RQ3 indicated

mixed results in the determination of whether teachers having been evaluated using KEEP affected their understanding among the performance ratings. The remaining 24 perceptions were not affected by whether the teacher had been evaluated using KEEP. On average, teachers who had been evaluated better understood the differences in the ratings of their performance on incorporating strategies for differentiation and scaffolding for all students, engaging in ongoing and purposeful professional learning connected to student learning, and reflecting on practice and actively seeks opportunities for improvement than teachers who had not been evaluated.

The results of the 27 hypothesis tests related to RQ4 indicated that administrators mostly or completely understood the differences among the performance ratings of highly effective, effective, developing, or ineffective. The analysis related to RQ5 to determine whether school level affected administrators' understanding could not be conducted due to sample size issues.

The results of the 27 hypothesis tests related to RQ6 indicated there were no differences between teachers' and administrators' understanding among the performance ratings of highly effective, effective, developing, or ineffective. The analysis related to RQ7 to determine whether school level affected the differences between teachers' and administrators' understanding could not be conducted due to sample size issues.

The results of hypothesis test related to RQ8 indicated that teachers mostly or completely understood the calculation of the final summative teacher evaluation rating. Additionally, the results of the hypothesis test related to RQ9 indicated that teachers' understanding of the final summative teacher evaluation rating was not affected by school level. However, the results of the hypothesis test related to RQ10 indicated that whether

teachers had been evaluated did affect their understanding of the final summative teacher evaluation rating. On average, teachers' understanding of the calculation of the final summative teacher evaluation rating was affected by whether the teacher has been evaluated using KEEP. Teachers who had been evaluated with the KEEP evaluation tool understood the calculation of the final summative rating better than those who had not yet been evaluated using the KEEP evaluation tool.

The results of hypothesis test related to RQ11 indicated that administrators mostly understood the calculation of the final summative teacher evaluation rating. The analysis related to RQ12 could not be conducted to determine whether school level affected administrators' understanding due to the sample size issues. The results of the hypothesis testing related to RQ13 indicated that on average, administrators understood the calculation of the final summative teacher evaluation rating better than teachers understood it. The analysis related to RQ14 could not be conducted to determine whether school level affected the differences between teachers' and administrators' understanding due to sample size issues.

The results of the four hypothesis tests related to RQ15 indicated that teachers agreed that the district was using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions. The results of the hypothesis tests related to RQ16 indicated that teachers' perceptions that the district was using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development were not affected by school level. However,

teachers' perceptions that the district was using the KEEP Model for personal decisions were affected by school level. Elementary teachers' perceptions were significantly lower than middle school teachers perceptions were. The results of the four hypotheses tests related to RQ17 indicated whether a teacher had been evaluated using KEEP did not affect their perceptions that the district was using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions.

The results of the four hypothesis tests related to RQ18 indicated that administrators agreed that the district was using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions. The analysis related to RQ19 could not be conducted to determine whether school level affected the differences between teachers' and administrators' understandings due to sample size issues.

The results of the four hypothesis tests related to RQ20 indicated that there were no differences between teachers' and administrators' perceptions that the district was using the KEEP Model for continual improvement of instruction, to meaningfully differentiate performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions. The analysis related to RQ21 could not be conducted to determine whether school level affected the differences between teachers' and administrators' understanding due to small sample sizes for administrators.

Findings Related to the Literature

This section contains a discussion of the outcomes of the current research study as they relate to the existing and pertinent literature identified in chapter two associated with perceptions of teacher evaluation models. A contrast of the results of the current research study to the existing literature discussed in chapter two yielded many similarities and differences. The findings related to the literature are presented below in the order of the research questions.

RQ1, RQ2, and RQ3 of this study focused on the extent that teachers understand the differences among the teacher performance ratings of highly effective, effective, developing, or ineffective. The findings indicated that the teachers mostly understood or completely understood the differences among the performance ratings of highly effective, effective, developing, or ineffective. The current finding is in contrast to the findings of Pizzi (2009), which highlighted teacher's perceptions of the "effectiveness of standards-based evaluation systems as a means of improving instruction and student achievement" (p. 3). The study was conducted at a large urban high school located in the northeastern United States. Pizzi found that a majority of the teachers were not familiar with the standards in the district evaluation tool. The results of RQ4 indicated that administrators mostly or completely understood the differences among the performance ratings of highly effective, effective, developing, or ineffective. These findings are in contrast to the findings of Summerville (2014) that administrators lack an understanding of statistical models used in the evaluation system.

Both teachers and administrators believed that the district was using the KEEP Model for continual improvement of instruction, to meaningfully differentiate

performance, to provide useful feedback that identifies needs and guides professional development, and to inform personnel decisions. These results are supportive of the findings of NCES (1994) that found that 63% of the teachers surveyed stated they could design a plan for their professional development based on their last performance evaluation, and 74% felt that their last evaluation had positively influenced their teaching skills. The current research study is also supportive of Ruffini et al.'s (2012) findings that teachers perceived the evaluation models was seen to be a strong method of supporting teachers' professional development because of its collaborative and reflective nature.

The current study findings are also in contrast to previous research that reports that evaluations are rarely reviewed by other leaders within the district, except in extremely rare cases when a supervisor rates a teacher as unsatisfactory (Marshall, 2005). The current findings are also in contrast to Summerville's (2014) study conducted at Vanderbilt University. Summerville found that many principals did not use data to make informed decisions for evaluations; 84% trusted the teacher observation to help make these decisions (Summerville, 2014). The findings of the current study are in contrast to Princess Towe's (2012) findings that teachers perceived the evaluation system being implemented to have little effect on improving the effectiveness of teachers and that the summative evaluation has the most effect on professional development goals. Also, in contrast to the current study are Kersten and Israel (2005) findings that administrators view current evaluation practices as not as effective due to the limited time administrators can spend on the process and the unclear goals or feedback provided by district procedures. Last, the findings of this study are in contrast to Barton (2010) who found

that administrators indicated that the teacher evaluation model lacked a clear purpose for professional development or staffing decisions.

Conclusions

The conclusion section contains the implications for action, which included how the results of this study can be applied to the field of education. Further research in the area of teacher and administrator perceptions of teacher evaluation models is also suggested. Last, concluding remarks are presented.

Implications for action. The results of the current research study provide implications for action for school administrators. Teacher evaluation is essential to the development of effective teachers within a district. District G can rely on the details provided within the current study to use the KEEP evaluation to ensure appropriate professional development and feedback that promote the success of every educator within the district. Specific recommendations for future action are warranted based on the results of this study.

The results of hypothesis testing related RQ1, RQ2, and RQ3 showed that teachers who had been evaluated better understood the differences in the ratings of their performance on incorporating strategies for differentiation and scaffolding for all students, engaging in ongoing and purposeful professional learning connected to student learning, and reflecting on practice and actively seeks opportunities for improvement than teachers who had not been evaluated. District G may be inclined to provide professional development for all teachers to identify objectives within the constructs of KEEP and apply them towards their yearly goals.

Based on the findings of RQ13, which indicated administrators understood the calculation of the final summative teacher evaluation rating better than teachers understood it, District G may benefit from facilitating professional development related to the differences in KEEP ratings and calculation of the summative rating for all teachers. All administrators and teachers at District G may also benefit from professional development focused on the appropriate plan for providing all teachers with an overview of the KEEP process yearly. School district leaders should be aware of their teacher's and administrator's perspective on the KEEP process and their level of familiarity on the ratings. It may also be appropriate for building leaders to administer surveys for teachers to self-assess themselves throughout the year. Administrators would benefit from collaborating during the summer to create consistency of "ratings" of teachers based on "mock" information.

Another recommendation is to add the component of reflection to administrators' comments and ratings. District leaders should provide adequate time throughout the year for building administrators to align their ratings and share feedback that is given to teachers. Reflection of the current completed KEEP evaluations could be vital in establishing accountability and ensure that building administrators are using the evaluation instrument for informed personnel decisions.

Recommendations for future research. The first recommendation is for the current research study to be replicated in each building during the 2016-2017 school year and to disaggregate the data by building. Building and district administrators can use the data to create a building specific professional development that aligns with the needs of the teachers. Teachers' self-assessment of their understanding of the ratings, summative

rating, and constructs could be a guide for each teacher as they develop their cycle goals. A second suggestion would be to replicate the current study in different school districts with similar and dissimilar characterizes to District G that have used the KEEP evaluation process. A replication of the study in similar school districts could allow district leaders to reflect on the professional development for their staff on the KEEP evaluation process and for comparisons to be made between school districts. Comparing the needs and recommendations from other districts may allow leaders to see various resources and needs that align with strengthening the instruction of their staff. Districts can also share resources for professional development, provide joint professional development, and compare results. Another study could be conducted when the component of reflection to administrators' comments and ratings is added to the evaluation. Replicating the study in a larger school district might also provide data about whether school level affects administrators' perceptions. A qualitative study could be conducted to get feedback that is more detailed from teachers and administrators on the benefits of the KEEP evaluation process and the professional development needed for both teachers and administrators. The last recommendation would be to replicate the study in two or three years to determine whether the results changed based upon the professional development provided.

Concluding remarks. The results of the present study can be contributed to the existing research relating to teachers and administrators perception on their own understanding of the calculation of the final summative rating and the six guidelines established within the Kansas evaluation process. Twenty-one research questions were identified to determine the perceptions of teachers and administrators on the KEEP

evaluation instrument within District G. Data were desegregated by school level, whether a teacher had been evaluated using KEEP, and by teachers and administrators. The results indicated that District G is in need of professional development for both teachers and administrators on the KEEP evaluation process. The results of this study should compel school districts and teachers to analyze their understanding of the KEEP instrument. While school districts strive to ensure highly effective educators in each classroom, it is imperative that the instrument used for evaluation integrates instructional practice and feedback that are supported by research.

References

Barton, S. N. B. (2010). *Principals' perceptions of teacher evaluation practices in an urban school district* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses. (UMI No. 3420551)

Bonavitacola, A. C. (2014) *Teachers' perceptions of the impact of the McREL teacher evaluation system on professional growth* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3623266)

Breedlove, P. H. (2011) *Teacher evaluation in North Carolina: Teacher perceptions during a time of change* (Doctoral dissertation). Retrieved from http://thescholarship.ecu.edu/handle/10342/3549

Burke, P., & Krey, R. (2005). *Supervision: A guide to instructional leadership* (2nd ed.). Springfield, IL: Charles C. Thomas Publisher.

Calabrese, R. L. (2006). *The elements of an effective dissertation and thesis: A step-by-step guide to getting it right the first time*. Lanham, MD: Rowman & Littlefield Education.

Cubberley, E. P. (1929). *Public school administration* (3rd ed.). Boston, MA: Houghton Mifflin.

Culbertson, J. (2012). Putting the value in teacher evaluation. *Phi Delta Kappan,94*(3), 14-18.

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

Danielson, C. (2002). Developing and retaining quality classroom teachers through mentoring: OCLC first search. *The Clearing House, 75*(4), 183-185.

Danielson, C. (2007). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.

Danielson, C. (2008). *The handbook for enhancing professional practice: Using the framework for teaching in your school.* Alexandria, VA: Association for Supervision and Curriculum Development.

Danielson, C. (2011). Evaluations that help teachers learn. *Educational Leadership, 68*(4), 35-39.

Danielson, C., & McGreal, T. (2000). *Teacher evaluation to enhance professional practice.* Alexandria, VA: Association for Supervision and Curriculum Development.

Darling-Hammond, L. (1999). *Professional development for teachers: Setting the stage for learning from teaching.* Santa Cruz, CA: Center for the Future of Teaching and Learning.

Darling-Hammond, L. (2012). *Creating a comprehensive system for evaluating and supporting effective teaching.* Stanford, CA: Stanford Center for Opportunity Policy in Education.

Doherty, J. F. (2009). *Perceptions of teachers and administrators in a Massachusetts suburban school district regarding the implementation of a standards-based teacher evaluation system* (Doctoral Dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI 3431766)

Dolph, D. A., & Kowalski, T. J. (2015). Principal Dispositions Regarding the Ohio Teacher Evaluation System. *AASA Journal of Scholarship & Practice*, *11*(4), 4-20.

Dyer, K. M., & Renn, M. T. (2010). Getting an education: School leaders need specialized development. *Leadership in Action, 29*(6), 3-7.

Frank, V. (2013). Evaluations serve as pathways for professional growth: Teacher-led teams help build evaluation system that promotes learning. *The Learning System, 8*(2), 4-5.

Garth-Young, B. (2007). *Teacher evaluation: Is it an effective practice? A survey of all junior high/middle school principals in the State of Illinois* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Data Base (UMI No. 304848276)

Glanz, J. (1991). *Bureaucracy and professionalism.* Cranbury, NJ: Associated University Presses.

Glowacki, H. (2013). *Illinois elementary principals' perceptions of the teacher evaluation process of special education teachers* (Doctoral dissertation). Retrieved from https://www.ideals.illinois.edu/bitstream/handle/2142/44303/Heather_Glowacki.pdf?sequence=1

Hopkins, G. (2001). The best and worst things about being a principal. *Education World.* Retrieved from http://www.educationworld.com/a_admin/admin/admin253.shtml

Hunter, M. (1980). Six types of supervisory conferences. *Educational Leadership, 37*(5), 408–412

Johnson, J., Dupuis, V., Musial, D., Hall, G., & Gollnick, D. (1995). *Introduction to the foundations of American education* (10th ed.). New York, NY: Simon & Shuster.

Kansas State Department of Education. (2014). *KEEP handbook: KEEP educator evaluations system guidelines*. Retrieved June 23, 2015, from http://community.ksde.org/Portals/44/Documents/0%20%20KEEP%20Handbook%20January%202014%20web.pdf

Kansas State Department of Education. (2015*). District report card*. Retrieved June 23, 2015, from https://online.ksde.org/rcard/district.aspx?org_no=D0231

Kersten, T. A., & Israel, M.S. (2005). Teacher evaluation: Principals' insights and suggestions for improvement. *Planning and Changing*, *36*(1-2), 47-67.

Larsen, M. (2004). A critical analysis of teacher evaluation policy trends. *Australian Journal of Education, 49*(2), 3.

Lunenburg, F. C., & Irby, B. J. (2008). *Writing a successful thesis or dissertation: Tips and strategies for students in the social and behavioral sciences*. Thousand Oaks, CA: Corwin Press.

Marshall, K. (2005). It's time to rethink teacher supervision and evaluation. *Phi Delta Kappan, 86*(10), 727–735.

Marzano, R. J. (2011). *The Marzano teacher evaluation model.* Englewood, CO: Marzano Research Laboratory.

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

Medley, D. M. (1977). *Teacher competence and teacher effectiveness: A review of process-product research*. Washington, DC: American Association of Colleges of Teacher Education.

McGreal, T. (1983). *Successful teacher evaluation.* Alexandria, VA: ASCD.

Morelock, M. (2008). *Investigating promising practice of teacher evaluation in two California charter schools* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3324990)

National Center for Educational Statistics. (1994). *Public elementary teachers' views on teacher performance evaluations.* Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

National Education Association. (2010). *Teacher assessment and evaluation: The national education association's framework for transforming education systems to support effective teaching and improve student learning.* Washington, DC: Author.

Nixon, A., Packard, A., & Dam, M. (2013). Principals judge teachers by their teaching. *Teacher Educator*, *48*(1), 58-72.

Pizzi, J. (2009). *Urban secondary teachers' perceptions of a standards-based teacher evaluation system* (Doctoral dissertation). Retrieved from eScholarship@BC.

Rindler, B. (1994). *The attributes of teacher evaluation systems that promote teacher growth as perceived by teachers of intensive English programs* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 9417S22)

Ruffini, S., Makkonen, R., Tejwani, J., & Diaz, M. (2014). *Principal and teacher perceptions of implementation of multiple-measure teacher evaluation systems in Arizona.* Washington, DC: U.S. Department of Education.

Sackett, P. R., & Larson, J. R. (1990). Research strategies and tactics in industrial and organizational psychology. *Handbook of industrial and organization psychology*, pp.419-489. Palo Alto, CA: Consulting Psychologist Press.

Sheppard, J. (2013). *Perceptions of teachers and administrators regarding the teacher evaluation process* (Doctoral dissertation). Paper 852. Retrieved from http://digitalcommons.georgiasouthern.edu/cgi/viewcontent.cgi?article=1856&context=etd

Sullivan, S., & Glanzer, J. (2000). *Supervision that improves teaching: Strategies and techniques.* Thousand Oaks, CA: Corwin.

Summerville, D. R. (2014). Study: Teacher data remain untapped. *Education Week, 34*(3), 1, 12-13.

Sutton, S. R. (2008) *Teachers' and administrators' perceptions of teacher evaluation* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 340518)

Tracy, S. J. (1995). How historical concepts of supervision relate to supervisory practices today. *Clearing House*, *68*, 320-325.

Weisberg, D., Sexton, S., Vlulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness.* Brooklyn, NY: The New Teacher Project.

White House. (2015) *Race to the Top.* Washington, DC: Author.

Wise, E., Darling-Hammond, L., McLaughlin, M., & Bernstein, H. (1984). *Teacher evaluation: A study of effective practices*. Santa Monica, CA: RAND.

Zimmerman, S., & Deckert-Pelton, M. (2003). *Evaluating the evaluators: Teachers' perceptions of the principal's role in professional evaluation.* NASSSP Bulletin, *87*(63), 28-37.

Appendices

Appendix A: Teacher Survey

Article I. KEEP Teacher Survey

The Kansas Educator Evaluation Protocol (KEEP) utilizes a four level rating rubric to evaluate teacher performance: Ineffective Developing Effective Highly Effective Please identify your understanding of the differences between these ratings for each of the following teaching behaviors included on the KEEP tool:

1. Planning and alignment of instruction to meet student learning needs and developmental levels*

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

2. Using a variety of teaching approaches and resources *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

3. Adapting instruction to meet student needs *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

4. Taking steps to gain knowledge of all students' individual differences *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

5. Using the knowledge of students to create a culture of respect among all students *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

6. Collaborating with students to promote student ownership of the learning *

- Do not Understand
- Somewhat Understand
- Mostly Understand

- Completely Understand

7. Establishing a safe, respectful, and academically challenging environment *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

8. Displaying knowledge of content by encouraging the use of multiple representations, explanations, and a wide variety of experiences *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

9. Using strategies to build student understanding of content *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

10. Using problem solving, critical thinking skills, and technology to explore and deliver content *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

11. Exploring and delivering content through real world applications of knowledge *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

12. Collaborating with colleagues to provide cross curricular learning opportunities for students *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

13. Planning rigorous and meaningful activities to meet the needs of all students *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

14. Using objectives that align with district, state, and/or national standards *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

15. Meeting the needs of all students *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

16. Providing opportunities for students to demonstrate learning through a variety of assessment methods *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

17. Using assessment data to inform instruction *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

18. Providing assessment feedback to promote students' responsibility for their own learning *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

19. Using a variety of strategies to engage and challenge students in a variety of learning situations *

- Do not Understand
- Somewhat Understand

- Mostly Understand
- Completely Understand

20. Incorporating strategies for differentiation and scaffolding for all students *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

21. Engaging students in higher order thinking *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

22. Engaging in ongoing and purposeful professional learning connected to student learning *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

23. Reflecting on practice and actively seeks opportunities for improvement *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

24. Analyzing and reflecting on student data to impact student growth *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

25. Collaborating with multiple stakeholders in school and professional activities *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

26. Using a variety of methods of communication with stakeholders *

- Do not Understand

- Somewhat Understand
- Mostly Understand
- Completely Understand

27. Demonstrating leadership skills used to support and improve student learning *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

Article II. The following statements are associated with the Kansas Educator Evaluation Protocol (KEEP) and the implementation of the new evaluation system within the school district. For each statement, please indicate your level of agreement.

28. The district is using the KEEP model for continual improvement of instruction. *

- Strongly Disagree
- Disagree
- Agree
- Strongly Agree

29. The district is using the KEEP model to meaningfully differentiate instruction. *

- Strongly Disagree
- Disagree
- Agree
- Strongly Agree

30. The district is using the KEEP model to identify teacher needs and guide professional development. *

- Strongly Disagree
- Disagree
- Agree
- Strongly Agree

31. The district is using the KEEP model to make informed personnel decisions. *

- Strongly Disagree
- Disagree

- Agree
- Strongly Agree

32. I understand the calculation of the summative rating. *

- Strongly Disagree
- Disagree
- Agree
- Strongly Agree

Article III. Please respond to the following demographic questions.

33. I currently work as a *

- Elementary Teacher (K-4)
- Middle School Teacher (5-8)
- High School Teacher (9-12)
- Teacher at Multiple School Levels
- Other:

34. I have been evaluated using KEEP during the following semesters (check all that apply): *

- I have not been evaluated using KEEP
- Fall 2014
- Spring 2015
- Fall 2015

Appendix B: Administrator Survey

Article IV. KEEP Administrator Survey

The Kansas Educator Evaluation Protocol (KEEP) utilizes a four level rating rubric to evaluate teacher performance: Ineffective Developing Effective Highly Effective Please identify your understanding of the differences between these rating for each of the following teaching behaviors included on the KEEP tool:

1. Planning and alignment of instruction to meet student learning needs and developmental levels*

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

2. Using a variety of teaching approaches and resources *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

3. Adapting instruction to meet student needs *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

4. Taking steps to gain knowledge of all students' individual differences *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

5. Using the knowledge of students to create a culture of respect among all students *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

6. Collaborating with students to promote student ownership of the learning *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

7. Establishing a safe, respectful, and academically challenging environment *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

8. Displaying knowledge of content by encouraging the use of multiple representations, explanations, and a wide variety of experiences *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

9. Using strategies to build student understanding of content *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

10. Using problem solving, critical thinking skills, and technology to explore and deliver content *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

11. Exploring and delivering content through real world applications of knowledge *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

12. Collaborating with colleagues to provide cross curricular learning opportunities for students *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

13. Planning rigorous and meaningful activities to meet the needs of all students *

- Do not Understand
- Somewhat Understand

- Mostly Understand
- Completely Understand

14. Using objectives that align with district, state, and/or national standards *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

15. Meeting the needs of all students *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

16. Providing opportunities for students to demonstrate learning through a variety of assessment methods *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

17. Using assessment data to inform instruction *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

18. Providing assessment feedback to promote students' responsibility for their own learning *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

19. Using a variety of strategies to engage and challenge students in a variety of learning situations *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

20. Incorporating strategies for differentiation and scaffolding for all students *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

21. Engaging students in higher order thinking *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

22. Engaging in ongoing and purposeful professional learning connected to student learning *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

23. Reflecting on practice and actively seeks opportunities for improvement *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

24. Analyzing and reflecting on student data to impact student growth *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

25. Collaborating with multiple stakeholders in school and professional activities *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

26. Using a variety of methods of communication with stakeholders *

- Do not Understand
- Somewhat Understand
- Mostly Understand

- Completely Understand

27. Demonstrating leadership skills used to support and improve student learning *

- Do not Understand
- Somewhat Understand
- Mostly Understand
- Completely Understand

Article V.The following statements are associated with the Kansas Educator Evaluation Protocol (KEEP) and the implementation of the new evaluation system within the school district. For each statement, please indicate your level of agreement.

28. The district is using the KEEP model for continual improvement of instruction. *

- Strongly Disagree
- Disagree
- Agree
- Strongly Agree

29. The district is using the KEEP model to meaningfully differentiate instruction. *

- Strongly Disagree
- Disagree
- Agree
- Strongly Agree

30. The district is using the KEEP model to identify teacher needs and guide professional development. *

- Strongly Disagree
- Disagree
- Agree
- Strongly Agree

31. The district is using the KEEP model to make informed personnel decisions. *

- Strongly Disagree
- Disagree
- Agree
- Strongly Agree

32. I understand the calculation of the summative rating. *

- Strongly Disagree
- Disagree
- Agree
- Strongly Agree

Article VI. Please respond to the following demographic questions.

33. I currently work as a *

- Central Office Administrator
- Elementary Administrator (K-4)
- Middle School Administrator (5-8)
- High school Administrator (9-12)

Appendix C: IRB

BAKER UNIVERSITY

SCHOOL OF EDUCATION
GRADUATE DEPARTMENT

Date:
IRB PROTOCOL NUMBER _____
(IRB USE ONLY)

IRB REQUEST
Proposal for Research
Submitted to the Baker University Institutional Review Board

I. **Research Investigator(s)** (Students must list faculty sponsor first)

Department(s) School of Education Graduate Department

Name	Signature	
1. Dr. Susan Rogers	Susan Rogers,	Major Advisor
2. Margaret Waterman	Margaret Waterman,	Research Analyst
3. Harold Frye	_____,	University Committee Member
4.		External Committee Member

Principal Investigator: Brooke L. Hughes Brooke L. Hughes
Phone: 913-485-3380
Email: Hughesb@usd231.com
Mailing address: 15911 W. 162nd Street, Olathe, KS 66062

Faculty sponsor: Dr. Susan Rogers
Phone: 913-344-1226 (Office) 785-230-2801 (Cell)
Email: srogers@bakeru.edu

Expected Category of Review: ___Exempt __X__ Expedited __Full

II: Protocol: (Type the title of your study)

Administrators' and Teachers' Perceptions of the Kansas Educator Evaluation Protocol

Summary

In a sentence or two, please describe the background and purpose of the research.

The study will be conducted within District G. District G is located in Johnson County, Kansas and is home to 5,452 students. Within District G, there are eleven schools consisting of seven elementary schools, three middles schools, and one high school. The data that will be gathered from this research could identify teachers' and administrators' perceptions of the KEEP, and whether they fully understand the calculation of the final summative rating and the six guidelines established within the Kansas evaluation process. The findings of the study may be used to improve the evaluation process in District G, feedback to teachers provided by administrators, and could add to the body of research related to teacher evaluation.

Briefly describe each condition or manipulation to be included within the study.

There will be no condition or manipulation in this study.

What measures or observations will be taken in the study? If any questionnaire or other instruments are used, provide a brief description and attach a copy.
Will the subjects encounter the risk of psychological, social, physical or legal risk? If so, please describe the nature of the risk and any measures designed to mitigate that risk.

Participants will be asked to complete a survey that includes Likert Scale items and demographic information (see attached surveys). One survey will be completed by administrators and one will be completed by teachers within District G. Subjects that are teachers within District G will be asked their perception of their own knowledge of the Kansas Educator Evaluation Program (KEEP). Subjects will be asked about their perception of their own growth through the evaluation instrument used within District G. Some demographic questions will be asked, including their school level taught, number of times evaluated using KEEP, and number of years they have taught.
Subject will not encounter the risk of psychological, social, physical or legal risk.
Subjects that are administrators within District G will be asked about their perception of their own knowledge of the Kansas Educator Evaluation Program (KEEP). Subjects will be asked about their perception of their own growth through the evaluation instrument used within District G. Some demographic questions will be asked, including their school level they serve, number of times evaluated using KEEP, and number of years within their current administrative role.

Will any stress to subjects be involved? If so, please describe.

Subjects will not be subjected to any stress.

Will the subjects be deceived or misled in any way? If so, include an outline or script of the debriefing.
Subjects of this study will not be deceived or mislead in any way.

Will there be a request for information which subjects might consider to be personal or sensitive? If so, please include a description.

Subjects that are teachers within District G will be asked their perception of their own knowledge of the Kansas Educator Evaluation Program (KEEP). Subjects will be asked about their perception of their own growth through the evaluation instrument used within District G. Some demographic questions will be asked, including their school level taught, number of times evaluated using KEEP, and number of years they have taught.

Subjects that are administrators within District G will be asked about their perception of their own knowledge of the Kansas Educator Evaluation Program (KEEP). Subjects will be asked about their perception of their own growth through the evaluation instrument used within District G. Some demographic questions will be asked, including their school level they serve, number of times evaluated using KEEP, and number of years within their current administrative role.

Will the subjects be presented with materials which might be considered to be offensive, threatening, or degrading? If so, please describe.

The subjects in this study will not be presented with materials that might be considered to be offensive, threatening, or degrading.

Approximately how much time will be demanded of each subject?

Participants will spend approximately ten to fifteen minutes completing the knowledge survey.

Who will be the subjects in this study? How will they be solicited or contacted? Provide an outline or script of the information which will be provided to subjects prior to their volunteering to participate. Include a copy of any written solicitation as well as an outline of any oral solicitation.

Subjects of this study will be teachers and administrators within District G that use the Kansas Educators Evaluation Protocol. Each subject will be contacted via email (see attached letters).

What steps will be taken to insure that each subject's participation is voluntary? What if any inducements will be offered to the subjects for their participation?

Completion of the survey is strictly voluntary, with completion of the survey indicating willingness to participate.

How will you insure that the subjects give their consent prior to participating? Will a written consent form be used? If so, include the form. If not, explain why not.

Voluntary completion of the survey will indicate consent by the subject. Subjects will be notified of implied consent in the initial email.

Will any aspect of the data be made a part of any permanent record that can be identified with the subject? If so, please explain the necessity.

Data gleaned from this study will not be made part of any permanent record.

Will the fact that a subject did or did not participate in a specific experiment or study be made part of any permanent record available to a supervisor, teacher or employer? If so, explain.

No data or participation will be made part of any permanent record available to a supervisor, teacher, or employer.

What steps will be taken to insure the confidentiality of the data? Where will it be stored? How long will it be stored? What will be done with it after the study is completed?

Survey results are confidential. Subjects will not be identified, recorded, or reported in the results of this study. Anonymous data will be stored in an external storage device and destroyed after 2 years.

If there are any risks involved in the study, are there any offsetting benefits that might accrue to either the subjects or society?

There are no risks associated with the participation of this study. Benefits of this study include advancement in the field of teacher evaluation for District G and continued tailored professional development opportunities for teachers and administrators.

Will any data from files or archival data be used? If so, please describe.

No archival data will be used in this study.

Appendix D: IRB Approval

Baker University Institutional Review Board

October 13, 2015

Dear Brooke Hughes and Dr. Rogers,

The Baker University IRB has reviewed your research project application and approved this project under Expedited Status Review. As described, the project complies with all the requirements and policies established by the University for protection of human subjects in research. Unless renewed, approval lapses one year after approval date.

Please be aware of the following:

1. Any significant change in the research protocol as described should be reviewed by this Committee prior to altering the project.
2. Notify the IRB about any new investigators not named in original application.
3. When signed consent documents are required, the primary investigator must retain the signed consent documents of the research activity.
4. If this is a funded project, keep a copy of this approval letter with your proposal/grant file.
5. If the results of the research are used to prepare papers for publication or oral presentation at professional conferences, manuscripts or abstracts are requested for IRB as part of the project record.

Please inform this Committee or myself when this project is terminated or completed. As noted above, you must also provide IRB with an annual status report and receive approval for maintaining your status. If you have any questions, please contact me at CTodden@BakerU.edu or 785.594.8440.

Sincerely,

Chris Todden EdD
Chair, Baker University IRB

Baker University IRB Committee
Verneda Edwards EdD
Sara Crump PhD
Erin Morris PhD
Scott Crenshaw

Appendix E: KEEP Administrator Survey Email Body

Your participation is completely voluntary, and responses will remain anonymous. Your name will not appear anywhere on the survey. You also have the option to not answer any question that may make you feel uncomfortable. Your answers will be combined with other participants and reported in summary form.

By completing and returning the survey, you are acknowledging that you are consenting to participate in this study. Your participation in the survey is extremely important for the completion of my dissertation. If you have any questions regarding this research, please contact me at hughesb@usd231.com.

Appendix F: KEEP Teacher Survey Email Body

Your participation is completely voluntary, and responses will remain anonymous. Your name will not appear anywhere on the survey. You also have the option to not answer any question that may make you feel uncomfortable. Your answers will be combined with other participants and reported in summary form.

By completing and returning the survey, you are acknowledging that you are consenting to participate in this study. Your participation in the survey is extremely important for the completion of my dissertation. If you have any questions regarding this research, please contact me at hughesb@usd231.com.

Appendix G: Results for RQ2, RQ3, and RQ6

Table G1

Hypothesis Test Statistics for RQ2

Hypotheses		*F*	*p*	*df1*	*df2*
H28	Alignment of Instruction	1.382	.254	2	169
H29	Variety of Teaching Approaches	1.550	.215	2	169
H30	Adapting Instruction	1.074	.344	2	167
H31	Knowledge of Differences	1.023	.362	2	169
H32	Culture of Respect	1.154	.318	2	169
H33	Collaborating with Students	.840	.434	2	169
H34	Establishing an Environment	.395	.674	2	169
H35	Multiple Representations	1.775	.173	2	169
H36	Strategies for Understanding	.779	.460	2	169
H37	Problem Solving and Tech.	1.646	.196	2	169
H38	Real World Applications	.496	.610	2	169
H39	Cross Curricular Learning	2.327	.101	2	169
H40	Rigorous Activities	.257	.774	2	169
H41	Alignment with Standards	.247	.782	2	169
H42	Meeting Student Needs	1.235	.293	2	169
H43	Assessment Methods	1.472	.232	2	169
H44	Assessment Data	1.523	.221	2	169
H45	Assessment Feedback	1.040	.356	2	169
H46	Strategies to Engage/Challenge	.100	.905	2	169
H47	Differentiation and Scaffolding	.786	.457	2	169
H48	Higher Order Thinking	1.609	.203	2	169
H49	Professional Learning	.630	.534	2	169
H50	Seeks Opportunities	.496	.610	2	169
H51	Impact of Data on Learning	.680	.508	2	169
H52	Collaborate with Stakeholders	1.040	.356	2	169
H53	Communication Methods	1.676	.190	2	169
H54	Leadership Skills	.710	.493	2	169

Table G2

Hypothesis Test Stastics for RQ3

Hypotheses		*t*	*df*	*p*
H55	Alignment of Instruction	.987	173	.325
H56	Variety of Teaching Approaches	1.118	173	.265
H57	Adapting Instruction	.277	171	.782
H58	Knowledge of Differences	1.235	173	.219
H59	Culture of Respect	1.876	173	.062
H60	Collaborating with Students	1.920	173	.056
H61	Establishing an Environment	.208	173	.836
H62	Multiple Representations	1.187	173	.237
H63	Strategies for Understanding	.759	173	.449
H64	Problem Solving and Tech.	.900	173	.369
H65	Real World Applications	.452	173	.652
H66	Cross Curricular Learning	1.219	173	.225
H67	Rigorous Activities	.467	173	.641
H68	Alignment with Standards	1.667	173	.097
H69	Meeting Student Needs	1.327	173	.186
H70	Assessment Methods	1.618	173	.108
H71	Assessment Data	.708	173	.480
H72	Assessment Feedback	.784	173	.434
H73	Strategies to Engage/Challenge	1.030	173	.304
H74	Differentiation and Scaffolding	2.697	173	*.008
H75	Higher Order Thinking	1.296	173	.197
H76	Professional Learning	2.035	173	*.043
H77	Seeks Opportunities	2.774	173	*.006
H78	Impact of Data on Learning	1.560	173	.121
H79	Collaborate with Stakeholders	1.465	173	.145
H80	Communication Methods	.671	173	.503
H81	Leadership Skills	1.314	173	.191

*$p < .05$

Table G3

Hypothesis Test Statistics for RQ6

Hypotheses		*t*	*df*	*p*
H136	Alignment of Instruction	.106	188	.915
H137	Variety of Teaching Approaches	-.239	188	.812
H138	Adapting Instruction	.685	186	.494
H139	Knowledge of Differences	.252	188	.801
H140	Culture of Respect	-.348	188	.728
H141	Collaborating with Students	-.641	188	.522
H142	Establishing an Environment	-.182	188	.856
H143	Multiple Representations	.653	188	.515
H144	Strategies for Understanding	.839	188	.403
H145	Problem Solving and Tech.	.071	188	.943
H146	Real World Applications	-.730	188	.467
H147	Cross Curricular Learning	-.384	188	.701
H148	Rigorous Activities	1.039	188	.300
H149	Alignment with Standards	.544	188	.587
H150	Meeting Student Needs	.451	188	.653
H151	Assessment Methods	.109	188	.913
H152	Assessment Data	.197	188	.844
H153	Assessment Feedback	.123	188	.902
H154	Strategies to Engage/Challenge	.145	188	.885
H155	Differentiation and Scaffolding	.182	188	.856
H156	Higher Order Thinking	.806	188	.421
H157	Professional Learning	.812	188	.418
H158	Seeks Opportunities	.052	188	.958
H159	Impact of Data on Learning	.393	188	.695
H160	Collaborate with Stakeholders	.427	188	.670
H161	Communication Methods	1.593	188	.113
H162	Leadership Skills	-.454	188	.650