TEACHERS’ DISPOSITIONS TOWARD THE OHIO TEACHER EVALUATION SYSTEM

Dissertation
Submitted to
The School of Education and Health Sciences of the UNIVERSITY OF DAYTON

In Partial Fulfillment of the Requirements for
The Degree of
Doctor of Philosophy in Educational Leadership

By
Beth A. Wyandt

UNIVERSITY OF DAYTON
Dayton, Ohio
May, 2016
TEACHERS’ DISPOSITIONS TOWARD THE OHIO TEACHER EVALUATION SYSTEM

Name: Wyandt, Beth Anne

APPROVED BY:

________________________
Theodore J. Kowalski, Ph.D.
Committee Chair

________________________
C. Daniel Raisch, Ph.D.
Committee Member

________________________
Carolyn S. Ridenour, Ed.D.
Committee Member

________________________
Teresa L. Thompson, Ph.D.
Committee Member

________________________
Kevin R. Kelly, Ph.D.
Dean, School of Education and Health Sciences
ABSTRACT

TEACHERS’ DISPOSITIONS TOWARD THE OHIO TEACHER EVALUATION SYSTEM

Name: Wyandt, Beth Anne
University of Dayton

Advisor: Dr. Theodore J. Kowalski

The Ohio Teacher Evaluation System (OTES) was created in response to the 2009 House Bill 1 mandate requiring the development of a state teacher performance assessment. This study examined K-12 public school teachers’ dispositions toward OTES after the first year of implementation. Data were collected from 142 teachers over a 4-week period in 2015 using a 17-item survey with Likert-type responses.

Findings revealed teachers’ overall dispositions were considerably more negative than positive; specifically 86% of the respondents had a negative or moderately negative disposition toward the state model. With respect to specific aspects of OTES, the three most positive dispositions concerned individualized staff development, principal compliance with OTES, and walk-through observations. The three most negative dispositions concerned the amount of time required of teachers, infusing student value-added scores, and infusing vendor assessment scores.
The levels of association between the criterion variable (a teacher’s overall disposition toward OTES) and each of three predictor variables (gender, years of teaching experience, and grade level assignment) were small and negative. Collectively, the three predictor variables accounted for only 4.5% of the variation in the criterion variable.

The findings have both professional and political implications. Most notably, negative dispositions reported in this study, especially those pertaining to the infusion of student assessments into teacher performance evaluations are highly controversial. Professionally, for example, many scholars, administrators, and teachers challenge the reliability, validity and fairness of using these measures. Politically, for example, policymakers should consider the negative dispositions in terms of improving existing policy. Accordingly, recommendations for improving OTES policy and for future research were made.
This dissertation is dedicated to all the educators who continually evaluate and reflect on their practice. No evaluation will ever capture the depth and breadth of your commitment to your students, nor the hours you spend making each learning experience memorable.

This dissertation is also dedicated to all the administrators who have worked tirelessly to support and encourage their teachers and students in the learning process. Thank you for your dedication, your leadership, and your service.
ACKNOWLEDGEMENTS

To my first teachers, my mother and father, my love and thanks for their unwavering belief in me. To my daughters, Nikki and Megan, and my grandsons, Christopher, Jonathan, and Samuel, for lifting me up with their prayers, words of encouragement and love.

To Dr. Theodore J. Kowalski, my admiration and appreciation are given to you for your sustained guidance during this endeavor. I was honored when you accepted me as your doctoral student. Without your direction, insight and expertise, this dissertation would not have been completed. Thank you to my committee members, especially Dr. Ridenour, for providing encouragement while challenging me to become a better researcher.

To my cohort, especially Dr. Lesley Evans and Dr. DeAnn Hurtado, thank you for the gentle nudges and supportive talks. I am indebted.

Finally, thank you to my best cheerleader and friend, my husband Roger, for the countless hours of encouragement and support. I could not have completed this dissertation without the sacrifices you made.
# TABLE OF CONTENTS

ABSTRACT ........................................................................................................ iv

DEDICATION .................................................................................................... vi

ACKNOWLEDGEMENTS ................................................................................ vii

LIST OF TABLES ............................................................................................. xi

CHAPTERS

I. INTRODUCTION .......................................................................................... 1
   Problem Statement ......................................................................................... 2
   Purpose and Research Questions ................................................................. 5
   Significance of the Study ............................................................................... 6
   Method .......................................................................................................... 6
      Study population ....................................................................................... 6
      Procedure .................................................................................................. 6
   Limitations .................................................................................................... 8
   Delimitations ................................................................................................. 9
   Definition of Terms ....................................................................................... 9
   Organization of the Study ........................................................................... 13

II. REVIEW OF RELATED LITERATURE ...................................................... 14
   Teacher Performance Evaluation ............................................................... 14
      History ...................................................................................................... 14
   Common Problems ....................................................................................... 18
      Neglecting formative evaluation ............................................................. 18
      Negative dispositions ............................................................................... 19
      Inadequate classroom observations and instrumentation .................. 21
Time restrictions ................................................................. 22
Evaluator bias and subjectivity ................................. 23
Lacking evaluator expertise ........................................ 25
Institutional restrictions .................................................. 26
The Ohio Teacher Evaluation System (OTES) ....................... 27
  History ........................................................................ 27
Evaluation components ............................................... 29
  New evaluation frameworks ........................................ 31
Shared responsibilities .................................................... 32
  Individualized staff development requirements .............. 32
Using Student Growth Measures to Evaluate Teacher
  Performance ............................................................... 34
  Value-added student learning data ......................... 35
  Student learning objectives .................................. 37
Linking Performance Evaluation to Individual Growth Plans .... 40
Summary ......................................................................... 45

III. METHODS ........................................................................ 47
  Research Questions ....................................................... 47
  Population ...................................................................... 47
  Instrumentation ............................................................. 48
    Reliability .................................................................. 49
    Validity ...................................................................... 50
  Data Collection ............................................................. 52
  Data Analysis ................................................................ 53
  Protection of Human Subjects ................................... 57

IV. REPORT OF FINDINGS ............................................................. 58
  Respondents’ profile ...................................................... 58
  Findings related to research questions .......................... 60
    Research question 1 .................................................... 60
    Research question 2 .................................................... 80
    Research question 3 .................................................... 81
    Research question 4 .................................................... 82
  Summary ....................................................................... 83
V. SUMMARY OF FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

Summary of Findings.............................................................................84
Demographic profile..............................................................................84
Findings pertaining to the research questions.................................85
Conclusions..........................................................................................87
Recommendations................................................................................91
Policy.....................................................................................................91
Future research....................................................................................91

REFERENCES..........................................................................................94

APPENDICES

Appendix A: Panel of Experts.................................................................111
Appendix B: Communication Sent to District Superintendents to Gain Permission to Contact Survey Participants........112
Appendix C: Communication Sent to Survey Participants.....................113
Appendix D: Teacher Dispositions toward OTES Survey.......................114
LIST OF TABLES

Table 1: Survey Instrument Questions Developed from OTES Elements……………...51
Table 2: Research Source of Survey Instrument Questions………………………...52
Table 3: Respondents’ Profile (N=142)………………………………………………...59
Table 4: Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions
   (N = 142) Statement 1………………………………………………………….61
Table 5: Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions
   (N = 142) Statement 2………………………………………………………….62
Table 6: Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions
   (N = 142) Statement 3………………………………………………………….63
Table 7: Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions
   (N = 142) Statement 4………………………………………………………….64
Table 8: Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions
   (N = 142) Statement 5………………………………………………………….65
Table 9: Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions
   (N = 142) Statement 6………………………………………………………….66
Table 10: Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions
   (N = 142) Statement 7………………………………………………………….67
Table 11: Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions
   (N = 142) Statement 8………………………………………………………….68
Table 12: Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142) Statement 9……………………………………………………………69

Table 13: Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142) Statement 10……………………………………………………………70

Table 14: Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142) Statement 11……………………………………………………………71

Table 15: Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142) Statement 12……………………………………………………………72

Table 16: Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142) Statement 13……………………………………………………………73

Table 17: Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142) Statement 14……………………………………………………………74

Table 18: Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142) Statement 15……………………………………………………………75

Table 19: Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142) Statement 16……………………………………………………………76

Table 20: Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142) Statement 17……………………………………………………………77

Table 21: Means and Disposition Classifications for Seventeen Specific Aspects of OTERs in Descending Order by Means (N = 142)……………………………78

Table 22: Overall Respondents’ Dispositions by Categories (N = 142)………………………80

Table 23: Correlation Coefficient and Levels of Association between Teachers’ Dispositions and the Three Predictor Variables………………………82
CHAPTER I
INTRODUCTION

One theme revealed consistently throughout teacher education literature is that the quality of classroom instruction is a critical factor in determining school effectiveness (National Council for Accreditation of Teacher Education, 2002; National Reading Panel, 2000; Sanders & Horn, 1998; Sanders, Horn, & Wright, 1997). Logically, both the education profession and many policymakers have deduced that improved teaching results in higher levels of student achievement. As a result, great attention has been given recently to mandating performance evaluation models that hold teachers accountable for student learning, especially by requiring the inclusion of value-added data into evaluation equations.

Over the past decade, state-mandated performance evaluation for teachers has changed. Expectedly, the administrative roles in this process have also changed. A principal’s responsibility, for example, has broadened especially in relation to formative evaluation, a process focused on continuous professional growth (Kowalski, Lasley, & Mahoney, 2008). Formatively, administrative supervision is intended to enhance teachers’ pedagogical skills with the ultimate goal of elevating student achievement (Marzano, Frontier, & Livingston, 2011).

As new evaluation models have been adopted, scholars have focused more intently on teacher dispositions than they did with previous models because findings
indicate attitudes affect performance measures and professional growth influences behavior in the classroom (Rushton, Morgan, & Richard, 2007). In fact, dispositions are considered just as critical to effective teaching as are skills and knowledge (Council of Chief State School Officer, 2011). Tschannen-Moran, Hoy, and Hoy (1998) argue that self-perceptions of teaching competence in the areas of skills, knowledge, strategies, or personality traits are judged against personal weaknesses or liabilities within the context of teaching. If a teacher does not feel that he or she is sufficiently competent, his or her effectiveness in the classroom is greatly reduced.

Extant literature on performance evaluation also supports the contention that the process, to be effective, should be a collaborative endeavor. In order for a collaborative process to be effective, teachers need to be receptive to the evaluative feedback they receive and they need to accept the premise that improving personal practice is a relevant professional norm. Concurrently, evaluators, such as principals, should have a mentoring relationship with teachers in order to establish trust and credibility. Research shows that teachers who feel supported are more apt to grow professionally (Dana & Yendel-Hoppy, 2009).

**Problem Statement**

The No Child Left Behind Act (NCLB) was promulgated to eliminate achievement gaps and to increase accountability for educating at-risk students (United States Department of Education, 2010). Initially, the NCLB set the goal that all students would be proficient in specified academic subjects by 2014. The reauthorization of the NCLB in 2010 included a revised blueprint for school reform. Specifically, lawmakers intended that the law would support systematic change to policies and practices ensuring
not just academic achievement, but growth for all learners. A major and controversial component of NCLB is measuring teacher effectiveness on the basis of student growth measures, such as value-added scores.

The Race to the Top initiative, a part of the American Recovery and Reinvestment Act of 2008, also affected school reform, especially by offering competitive grants to states that are willing to address the following four areas:

- Improvement of standards and assessments
- Adoption of better systems for communicating student progress
- Supportive efforts for elevating educator effectiveness
- Increased emphasis on and resources for rigorous interventions, particularly in the lowest-performing schools (USDE, 2014).

Moreover, the Race to the Top program has augmented NCLB with respect to encouraging and helping states promulgate policies mandating more rigorous teacher evaluation models. In Ohio, these two federal initiatives have contributed to a product know as the Ohio Teacher Evaluation System (OTES).

In order to understand the challenges OTES has presented to school principals, one needs to compare traditional norms for performance evaluation and staff development with new requirements common in state-mandated models. In the past, a teacher evaluation process usually was determined at the district level and as a result, purpose, processes, and criteria varied considerably across states and even among districts in the same state. Despite variations in rigor and methods, teacher evaluation commonly focused solely on summative assessments; the only purpose was to determine if a teacher’s performance was satisfactory (Youngs, 2013).
The traditional approach for staff development was generic rather than individualized (Zepeda, 2012). Commonly, all teachers in a district received the same experiences, delivered in sporadic one-day sessions. Such programs typically were determined by administrators without teacher input, were not job-embedded, were unsystematic, and were only marginally related to instruction (Darling-Hammond, 2004; Fullan, 2001; Killion, 2002).

Given their limited scope, neither traditional performance evaluation nor traditional staff development has been highly effective (Darling-Hammond, 2004; Fullan, 2001, Killion, 2002). As a result, many teachers have been dismissive of both performance evaluation and professional development, largely because the processes were unrelated to individual needs (DuFour & Marzano, 2011).

In the case of OTES, aspects of performance evaluation and professional development have become controversial among educators. The first such aspect is an OTES requirement that a percentage of a teacher’s annual evaluation must be based on student growth measures; the second is a requirement that principals must assist teachers to develop individual improvement plans annually. After the initial pilot year, teachers’ unions sought to change the percentage assigned to student growth measures and to expand the type of growth measures allowed. Policymakers did make the proposed changes in response to pressure from the teachers’ unions, but without substantiated evidence of teacher dispositions toward OTES. The implementation of this common statewide system lacks teacher involvement in the design and implementation of the evaluation process.
A critical need exists for teachers to be involved in planning the evaluation process (Clark, 1996). When teachers are involved concerns are diminished (Clark, 1996). Likewise, teachers who feel safe in the evaluation process are more apt to use evaluation outcomes to improve instructional practice (Tornero & Taut, 2010). Furthermore, Tornero and Taut (2010) suggest teacher perceptions should be considered when developing accountability policies.

Unfortunately, an empirical database detailing teacher dispositions on OTES has yet to be developed. The information void is problematic for policymakers who have the responsibility to determine the future of OTES and for administrators who have the responsibility to implement and improve OTES. Without information about teacher dispositions, future policy and administrative decisions are more likely to be made politically and subjectively and the potential for OTES to improve instructional practice restricted.

**Purpose and Research Questions**

This study examined teacher dispositions toward OTES, especially two provisions: basing performance evaluations partially on student growth measures and developing individual teacher growth plans annually. Data were collected from a defined population consisting of all public school teachers, Kindergarten through 12th grade in 8 districts and 34 schools located in Miami County, Ohio. The study was guided by four research questions:

1. What were the respondents’ dispositions toward specific aspects of OTES?
2. What were the individual respondents’ overall dispositions toward OTES?
3. What was the level of association between the criterion variable (teacher overall personal dispositions toward OTES) and each of three predictor variables (gender, years of teaching experience, and grade level teaching assignment)?

4. To what extent did the three predictor variables collectively account for the variability in the criterion variable?

**Significance of the Study**

The findings and conclusions reported in this study have value for several reasons. First, they contribute to the knowledge base of teacher performance evaluation and individual growth plans. Second, they provide data for principals who are responsible for evaluating teacher performance and for assisting teachers to develop their annual growth plans. Third, outcomes serve to inform future decisions made by policymakers and administrators.

**Method**

**Study population.** The defined study population (N=895) included all teachers in Miami County, Ohio who were employed during the 2014-15 school term and who were evaluated via OTES. Population members were identified by data available on the Ohio Department of Education’s website.

**Procedure.** Data were collected using a survey instrument, *Teacher Dispositions toward OTES*, developed by the researcher. The use of a survey method allowed the researcher to economically collect a large number of responses, protecting the anonymity of the respondents (Krathwohl, 2009).

Face and content validity of the instrument were determined by a panel of experts as recommended by Krathwohl (2009). The panel members are listed in Appendix A.
The panel assessed the appearance of the instrument and the instruments’ representation of the topic (Burton & Mazerolle, 2011). The survey was also reviewed by the author’s dissertation committee. The committee recommended using a mix of positively-keyed and negatively-keyed statements. Questions were rephrased accordingly. The committee also recommended separating the 17 items into two sections (1) Elements of OTES and (2) Experiences with OTES.

The instrument was field tested with 20 teachers outside the defined population. Based on their feedback, the researcher determined that additional changes were unnecessary. Cronbach’s alpha was used to determine the internal consistency of the survey as a measure of reliability.

The superintendent of each school district in Miami County was contacted via email to gain permission to email the survey to all K-12 teachers in his or her district (See Appendix B). One week after the initial email, a personal phone call was made to superintendents who had not responded to the email. Once permission was obtained, participants were contacted via email to participate in the survey (Appendix C). Individual email addresses were obtained from school district’s websites. A follow-up email was sent to non-responders after 2 weeks. The survey was administered using Survey Monkey Online System (See Appendix D).

Responses were electronically recorded and transferred to the Statistical Package for Social Sciences software (SPSS) for analysis by the researcher. Descriptive statistics were applied (means, standard deviations, frequency counts, and percentages) to answer the first research question. A four category typology was created to determine an individual respondents overall disposition toward OTES to answer the second research
question. A Pearson Product Moment Correlation was calculated to determine the level of association between teacher dispositions toward OTES and each demographic variable listed in the third research question. The coefficients were applied as a descriptive statistic as specified by Cohen, Cohen, West, and Aiken (2003). A multiple correlation coefficient (R) was calculated to determine the level of association between the criterion variable and the predictor variables collectively. The coefficient was then squared to determine the coefficient of determination (R²), a statistic that identifies the proportion of common variation attributable to the three predictor variables.

Anonymity was ensured through the use of the Survey Monkey instrument. A summary of findings was sent to the district and school administrators having supervisory responsibilities for the teachers comprising the study population. Results were also shared with responders who requested the information.

Limitations

There are four notable limitations to this study. The first is the size of the population. The study population consisted of 895 teacher employed in 8 districts all located in the same county. Thus, results should not be generalized to other populations. Secondly, the return rate is a limitation. Although 144 surveys were returned, 2 included no answers. Therefore, the number of surveys analyzed was 142 (N = 142) for a return rate of 15.8%. Therefore, external validity is limited. Thus findings are reflective of the respondents and not necessarily the entire study population. Third, inferences regarding non-responders should not be made. For inferences to be made regarding non-responders, information would have to be gathered. Finally, instruments that utilize self-report data also possess limitations since “…the researcher can never be sure that
individuals are expressing their true attitudes, interest, values, or personalities” (Gay, Mill, & Airasian, 2009, p. 153). Furthermore, studies that use self-report instruments share the common problem of the existence of a response set, an individual’s tendency to respond in a particular way (Gay, Mills, & Airasian, 2009).

**Delimitations**

The nature of the survey is a delimitation because there were no open-ended questions allowing respondents to clarify their responses. A Likert-type survey was chosen instead of utilizing open-ended questions. Administering the survey using an online measure instead of face-to-face administration also is a delimitation. Typically, return rates for the former are lower than the latter (Nulty, 2008).

**Definition of Terms**

*Classroom teacher* – A classroom teacher is an instructor responsible for a specific subject or grade level.

*Dispositions* – Dispositions are values, commitments, and professional ethics that influence behaviors toward students, families, colleagues, and communities and affect student learning, motivation, and development as well as the educator’s own professional growth. Dispositions are guided by beliefs and attitudes related to values such as caring, fairness, honesty, responsibility, and social justice. For example, they might include a belief that all students can learn, a vision of high and challenging standards, or a commitment to a safe and supportive learning environment (NCATE, 2014).

*Evaluator* – An OTES credentialed evaluator has completed state-sponsored training, passed an online assessment, is properly certified and approved by the local
board of education (NIET, 2012). Principals, assistant principals, and other district administration meeting these criteria can be an OTES evaluator.

*Formal Observation* – A formal observation is performance review conducted by an evaluator who spends a minimum of 30 minutes observing and recording data according to the Teacher Performance Evaluation Rubric in order to evaluate the level that best describes the teacher’s instruction (NIET, 2012).

*Formative Evaluation* – Formative evaluation is an ongoing process used to assist the teacher in improving their instructional effectiveness through the use of targeted, continuous feedback (Zepeda, 2012).

*Individual Professional Growth Plan* – An individual growth plan is developed by an individual teacher to focus on areas of improvement identified through evaluator feedback, student growth measures and the teacher’s self-assessment. The plan, developed for one academic year, utilizes specific professional development in order to improve teacher practice (NIET, 2012).

*Informal Observation (Walk-through)* – An informal observation is a tool to inform evaluation that provides the opportunity to gather evidence of instruction over a series of short classroom visits while giving targeted evidence-based feedback to teachers (NIET, 2012). It allows principals to visit classrooms more frequently and more purposefully.

*Individual Respondents’ Overall Dispositions toward OTES* – The general overall disposition toward OTES of a teacher in this study. Each teacher was categorized as having one of four dispositions as a result of his/her total score on the study survey (the sum of responses to the 17 statements each rated from 1 to 5) according to the following:
Negative disposition = 17-34 points; Moderately negative disposition = 22-43 points; Moderately positive disposition = 44-64 points; or Positive disposition = 65-85 points.

*Multiple Measures* – Multiple measures are factors in the teacher evaluation framework, which include: observation, walk-through, and student growth measures.

*Negatively-keyed statements* – Negatively-keyed statements are worded so that agreement with the statement indicates a more negative disposition toward the element or experience with OTES contained in the survey statement.

*Ohio Teacher Evaluation System (OTES)* – OTES is Ohio’s evaluation system for teachers based on teacher performance as measured by the Teacher Performance Rubric and by student performance as measured by student achievement data (value-added data, vendor assessments or student learning objectives).

*Performance Evaluation* – Performance evaluation describes the overall performance level of a teacher as measured by the teacher performance rubric during formal observations, informal observations (walkthroughs), collection of evidence and information conversations (NIET, 2012).

*Positively-keyed statements* – Positively-keyed statements are worded so that agreement with the statement indicates a more positive disposition toward the element or experience with OTES contained in the survey statement.

*Professional Development* – Professional Development is a continuous endeavor by a professional to increase the knowledge of their craft through the process of collaboration, reflection, teaching, and learning (Danielson & McGreal, 2000).
Public School – A public school in the United States is supported by public funds, free to all children between the ages of six and eighteen, and is governed by an elected board of education (Ohio State Bar Association, 2015).

Respondents’ Dispositions toward Specific Aspects of OTES – The overall response on the dispositions for a specific item on the Teachers’ Dispositions toward OTES Survey, e.g., individualized staff development is more effective than school-wide development, as determined by first assigning a value to each response and then calculating a mean score for each response item.

Student Learning Objective (SLO) – SLOs are identified goals formulated by a teacher or group of teachers as a means to measure progress towards a learning outcome or growth target for a group of students over a designated period of time (NIET, 2012).

Student Growth Measures – Student growth is the change in the individual, student’s achievement scores between two or more points in time (NIET, 2012).

Summative Evaluation – A summative evaluation is the rating given at the end of the evaluation cycle to describe the level of achievement a teacher attained. This evaluation becomes part of the teacher’s permanent record (NIET, 2012).

Teacher Category for Student Growth Measures – Teachers are assigned to a category for student growth measures based on the type of data used to measure student growth. There are four categories of teachers: “A” teachers have only value-added data to measure student growth; “A-1” teachers have value-added data plus another measure of student growth (vendor assessment or student learning objectives); “B” teachers have a combination of vendor assessments and student learning objectives to measure student
growth; and “C” teachers have only student learning objectives to measure student growth (NIET, 2012).

*Teacher Performance Evaluation Rubric* – The Teacher Performance Evaluation Rubric is scored holistically and provides a means to describe the overall level of teacher performance (NIET, 2012).

*Value-Added* – Value-added analysis is a statistical method that helps educators to measure the impact schools and teachers have on students’ academic progress rates from year to year (Ohio Department of Education, 2014).

*Vendor Assessment* – A vendor assessment is an approved assessment created by a national testing vendor to be utilized in calculating student growth.

**Organization of the Study**

Chapter 2 includes a review of the relevant literature on teacher evaluation performance, the Ohio Teacher Evaluation System, evidence of student performance, and individual growth plans. Chapter 3 describes the research methods and procedures utilized in developing and administering the survey instrument to complete the study. Chapter 4 presents an analysis of the research findings. Chapter 5 includes a summary of the findings, the conclusions drawn from the findings and implications for further research.
CHAPTER II
REVIEW OF RELATED LITERATURE

Teacher performance evaluation and professional development are essential components employed by public school districts to increase teachers’ knowledge and improve student learning. A review of the literature in this chapter was conducted on performance evaluation, student growth evidence, professional growth plans, and the Ohio Teacher Evaluation System (OTES).

The first section of this chapter addresses teacher performance evaluation. Included are a brief history of the processes, practices, and problems. The second section describes the Ohio Teacher Evaluation System (OTES), the program mandates, and specific aspects of teacher evaluation and individual teacher growth plans central to this study. The third section summarizes empirical studies on the use of student growth evidence in evaluation. The fourth section describes the nexus between performance evaluation and professional growth plans. The final section is a brief summary of the chapter.

Teacher Performance Evaluation

History. Education was not considered a discipline or field of study until the early 1800s. In fact, before that time, teachers were considered community public servants. Public schools did not have a board of education or district-level administrators. Local governments relied on clergy or knowledgeable and experienced
men to hire and evaluate teachers (Peterson, 1982). Often teacher evaluations were based primarily on public opinion or perception (Marzano, Frontier, & Livingston, 2011) and religious ideology and cultural norms typically were the criteria for making termination decisions (Peterson, 1982).

By the mid 1800s, with the rise of industry and a movement towards common schooling, teaching had become increasingly complex. Some teachers were hired based on expertise and certain pedagogical skills were considered necessary to be an effective teacher (Marzano et al., 2011). As a result, teacher performance evaluations became more focused on pedagogy rather than religious and community values.

John Franklin Bobbitt wrote a book in 1918 called The Curriculum that became the foundation for curricular design. Based on the standards of scientific management as developed by Frederick Taylor, the depiction of the school was a factory with a hierarchy comprised of the principal as manager, the teacher as supervisor, and the students as workers. Teachers were to be evaluated based on the measurement of specific work.

After World War II, emphasis was placed on “assisting the teacher to develop his or her unique skills” and “also tending to his or her emotional needs” (Marzano et al., 2011, p. 16). Evaluations continued to evolve and by 1980 almost 90% of teachers were evaluated using a clinical supervision model (Marzano et al., 2011). In the context of this paradigm, teachers had a conference with the principal prior to being observed in the classroom. Observation data were then analyzed, and the teacher had a post-observation conference with the evaluator (Goldhammer, 1969). Typically, formative evaluation was not a part of the process.
During the 1980s, Madeline Hunter’s model for teacher evaluation, based on seven specific steps, became normative (Fehr, 2001). The steps link teacher effectiveness with student growth; therefore, performance evaluations focused on helping teachers self-identify and analyze instructional behaviors, especially least effective practices and professional development needs.

Current literature indicates the purpose of teacher performance evaluation to provide feedback and guidance for improving professional practice while documenting accountability (Danielson & McGreal, 2000). In the book, *Enhancing Professional Practice: A Framework for Teaching*, Danielson (2007) advocates a model for a uniform but individualized teacher evaluation system. Her reasoning is two-fold: there needs to be a clear definition of the “what” of teaching and instruments and procedures providing evidence of the “how” of teaching. The model includes 76 criteria. Collectively, they provide data that culminates in one of four evaluations: unsatisfactory, basic, proficient, and distinguished. Based on the evaluation, a personal improvement plan is to be developed.

More recently, using student achievement metrics to establish teacher accountability has been proposed, primarily because some empirical studies have reported a nexus between effective teaching and student learning outcomes (Tucker & Strong, 2005). The No Child Left Behind Act (NCLB) and the more recent Race to the Top, for example, stress the importance of evaluating teachers based on student value-added metrics (Konstanopoulos, 2014).

Increasingly, students are given a voice in teacher performance evaluations through the use of surveys. The Measures of Effective Teaching (MET) project
evaluated the practice of using student surveys in over 2,000 classrooms and found that “the average student knows effective teaching when he or she experiences it” (Bill & Melinda Gates Foundation, 2010). Students were capable of differentiating between a teacher they merely like and one that is an effective instructor. The student evaluation contained discrete, observable elements such as classroom management and challenging instructional practices (Advocates for Children of New York, 2012). Currently, several states require or are piloting student surveys as an evaluation component: Alaska, California, Colorado, Connecticut, Georgia, Hawaii, Illinois, Massachusetts, North Carolina, Pennsylvania, Tennessee, and Utah (Advocates for Children of New York, 2012).

Likewise, parent surveys have garnered attention in the past few years with the use of the Danielson model and several similar state models. The models include a professional responsibility component of which communication to parents is an indicator. Originally, due to political and logistical implications, researchers discouraged the use of parent surveys in teacher evaluations (Loup, Garland, Ellett, & Rugutt, 1996; McGreal, 1983). Current research indicates involved parents have a positive impact on student performance, behavior, and graduation rates (Harvard Family Research Project, 2006). Peterson, Wahlquist, Brown, and Mukhopadhyso (2003) conducted a survey of 201 classrooms and found high ratings by parents combined with several other positive ratings proved to be a good indicator of effective teaching. Parents view their input as an important component of teacher performance evaluation (Advocates for Children of New York, 2012).
In summary, the process of teacher evaluation has evolved, and the contemporary norm calls for summative and formative assessments. The former determines if a teacher’s performance has been acceptable, the latter determines what should be done to improve teaching (Hughes, 2006; McGaghie, 1991). In the context of dual-purpose evaluation, teachers should be given the opportunities to reflect on personal performance so they are able to develop individualized growth plans.

**Common Problems**

**Neglecting formative evaluation.** Teacher evaluations may serve two purposes: measuring teacher performance and fostering professional growth (Boyd, 1989). The latter should give teachers useful feedback, the opportunity to learn new teaching techniques and counsel from principals and other teachers on how to improve instruction. To achieve these goals, evaluators set specific standards. According to Boyd (1989), the standards should (a) relate to important teaching skills, (b) be objective, (c) be clearly communicated to the teacher before assessments begin and reviewed after they are completed, and (d) be linked to the teacher’s professional development.

Alluding to formative evaluation, Darling-Hammond (2012) wrote, “A high-quality system should create a coherent, well-grounded approach to developing teaching” (p. 5). This type of evaluation samples the process of learning or improvement and focuses on how the recipient can continue to increase effectiveness. Thus, a formative evaluation provides feedback and recommendations for professional growth (McDougall, 2001; Manning, 1988). Many teachers, however, do not receive such feedback. For example 28.8% of San Bernardino, CA, teachers in 2002-03 and 32% of those in 2003-04 reported that they had received neither feedback nor suggestions during the evaluation.
process, which may explain why only 26% of teachers in one study reported that evaluation was useful and effective (Kimball & Milankowski, 2009).

Eighty-nine percent of respondents in a survey conducted by Zimmerman and Deckert-Pelton (2003) expressed the need for interaction with constructive feedback from the evaluator. Yet, only 11% of educators in this same study reported they had received constructive feedback.

Discussing the omission of formative evaluation, Manning (1988) wrote, “A chief reason that obsolete methods of teacher evaluation remain in use is that the purpose of evaluation is confused” (p. 1). Traditional purposes have included making tenure decisions, ensuring accountability, removing incompetent teachers, and awarding promotions. Often educators and legislators have been at odds regarding the purpose of performance evaluation. Educators have had a proclivity to view the process in the context of professional performance and growth, while legislators have had the proclivity to view the process in the context of quality assurance (Danielson & McGreal, 2000; Tucker & Strong, 2005).

Formative evaluation entails more than a measure of quality control by focusing on instructional improvement. Formative evaluation also provides an opportunity to create an individualized professional growth plan. When there is no formative component in performance evaluation teachers are not apt to develop personal learning goals, compare their level of performance to a desired level, and engage in effective actions to reduce the gap (Sadler, 1989).

**Negative dispositions.** According to the National Council for Accreditation of Teacher Education, (2014), dispositions are especially relevant because they represent the
values and commitments that define a teacher’s performance. Research indicates that teacher dispositions toward performance evaluation have been more negative than positive. A national study conducted by Duffett, Farkas, Rotherham, and Silva (2008), for example, reported that only 26% of the respondents viewed the process as effective and useful. Another study (Louis et al., 2010) reported that only 38% of teachers considered classroom observations helpful in relation to improving instruction.

Teachers’ dispositions toward performance evaluation and evaluators are critical largely because evaluation is a high-stakes, high-anxiety activity (Zimmerman & Deckert-Pelton, 2003). In order to accept the results of a performance evaluation, a teacher needs to believe the evaluation process is fair and accurate and the evaluator is credible.

Research reveals that many teachers are skeptical about the value of performance evaluation, especially in relation to the amount of time and work required of them. As examples, Calabrese, Sherwood, Fast, and Womack (2004) and Kimball (2003) reported that most teachers viewed performance evaluation as being unduly burdensome and as having little value. Moreover, studies by Natriello (1983) and Newton and Braithwaite (1988) found that many teachers considered performance evaluation to be of little benefit to either teachers or students. A study by Towe (2012) of four urban high schools found that teachers felt their evaluation had little or no influence on improving their performance.

Low self-efficacy has been cited as a concern with performance evaluations. Based on the findings in his study, Kimball (2003) proposes that self-efficacy influences how teachers feel about performance evaluation and deal with feedback they are given.
Those with low self-efficacy are more apt to devalue the evaluation process and to ignore improvement suggestions.

A number of studies (Oppenheim, 1994; Peterson, 2000; Spokane Education Association, 2011) have revealed teacher concerns about the ability of principals to conduct valid and reliable assessments. This belief is most prevalent when the evaluator lacks expertise in a teacher’s subject area or grade level (Oppenheim, 1994). Peterson (2000) found in a qualitative review that teachers have little confidence in teacher evaluations. Surveying 896 educators, the Spokane Education Association found that nearly 25% of the teacher members doubted that principals could conduct fair assessments (2011).

Studies often conclude that negative dispositions stem from experiences with ineffective evaluations. As reported in a study by Kersten and Israel (2005), 87% of administrators identified one or more barriers to effective evaluation, which lessened the impact on improved instruction. Himmelein (2009) found that 51% of principals she surveyed saw little connection between the evaluation process they were using and the improvement of teacher effectiveness.

**Inadequate classroom observations and instrumentation.** Classroom observations have been the primary data assessment technique for evaluating teachers over the past 100 years (Peterson, 2000). According to the Educational Research Service (2010), 99.8% of public school administrators conduct them. Yet, the literature reveals several concerns about instrumentation (Marshall, 2005) and techniques (Williams, Cray, Miller, & Protheroe, 2002). According to Williams and associates (2002), teachers often view observations as being “judgmental, intimidating, unfair, and damaging to
professional dialogue” (p. 18). The authors commented that negative judgments may have stemmed from the manner in which observations commonly have been conducted.

Analyzing poor observation techniques, Marzano (2007) posits that many principals commit measurement and sampling errors. The former occur when the evaluator does not adequately understand the assessment system and the latter occur when the ratings are not based on actual behavior. Although some level of error is inevitable, he recommends that principals should (a) require teachers to do self-assessments, (b) conduct unannounced observations, (c) record teaching behavior on video, and (d) allow teachers to challenge observation assessments (Marzano, 2012).

Assessment instruments may fail to address teachers’ content knowledge, pedagogy, and the effects on student learning. As such, they tend to emphasize easily measured, but not necessarily important instructional behavior. Youngs (2013) contends that observation instruments often quantify generic behavior that may or may not have a direct bearing on student learning and some evaluation instruments are merely checklists that administrators complete during or after a brief observation period. Such instruments do not require principals to explain their ratings and as a result, they have little or no value in terms of formative evaluation.

**Time restrictions.** The amount of time principals must devote to conducting performance evaluations is a frequently cited problem. Painter (2000), for example, surveyed 37 principals and 56.8% identified insufficient time as a barrier to conducting effective evaluations. Halverson, Kelley, and Kimball (2004), for instance, reported that as much as 25% of a principal’s workday is spent on evaluation activities related to evaluating employees. Kersten and Israel (2005), found that 47% of principals identified
time-related limitations to be a problem. For example, half of the principals indicated they devoted more than 10 hours to evaluating each non-tenured teacher. In another study by Sporte, Stevens, Healey, Jiang, and Hart (2013) principals reported spending more than 6 hours per teacher on the evaluation process (pre-conference, observation, write-up, and post conference) with the typical elementary school principal spending almost 2 weeks or 120 hours on observations (Sporte et al., 2013). In a study by Kimball (2003), principals reported they were often unable to complete evaluations or needed to cut back on the amount of evidence required in order to submit evaluations on time.

Evaluator bias and subjectivity. Bias and subjectivity contribute to inaccurate performance evaluations especially when conducted by a single person. Goldstein (2007) compared principal evaluations of teachers and master teacher evaluations of their peers. He found a higher level of accountability and decision-making in the evaluations conducted by master teachers. Moreover, teachers in this study preferred peer evaluations to principal evaluations. Having different people observe and evaluate teachers is thought to increase consistency and decrease evaluator bias (Sawchuck, 2013).

Principals report it is difficult to separate teacher contributions to the school organization from classroom performance (Scriven, 1981). Thus, both teachers and principals believe their evaluations are reflective of more than the criteria listed on the evaluation instrument (Orphonus, 2014). For example, teachers that have few parent complaints, volunteer for extra duty activities, or contribute to the school organization, often receive higher ratings even if their performance in the classroom is not reflective of the ratings. Likewise, ability, collegiality, and student satisfaction contribute to the overall ratings given by a principal (Jacob & Lefgren, 2006). Citing the halo effect,
Tucker and Strong (2005) contend that evaluator personal opinions and biases contaminate the evaluation process thus undermining credibility and trust conditions necessary for meaningful dialogue between administrators and teachers.

Biased ratings are more likely if the principal has had little or no training for conducting performance evaluations. According to Orphanos (2014), research documents the presence of gender, race, and age bias in some teacher evaluations. For instance, older teachers have been found to receive higher ratings than younger teachers (Ostrander, 1996).

Rinehart and Young (1996) found two types of bias in teacher evaluations. The first was gender, that is female teachers were rated higher than male teachers. The second was administrator perception of effective teaching; that is teachers were “under or overrated because they did not match the situational or dispositional category of the principal” (Rinehart & Young, 1996, p 321). Teachers were rated on their professionalism instead of their instructional practice.

In a more recent study examining the dual role of administrators as both evaluators and instructional coaches, Sporte et al. (2013) reported several teachers of the 31 teachers interviewed felt evaluators purposefully lowered their performance rating when teachers attempted to gain support for areas of weakness. For example, one teacher reported going to her principal for help with classroom management, but was penalized with a lower rating in this area on her performance evaluation. Donaldson (2012) examined 72 teachers’ perspectives on the evaluation process reporting approximately 50% of the respondents felt the evaluation process was unfair due to administrator bias. The ratings were influenced by the evaluators’ feelings about the teachers they assessed.
After surveying 896 teachers, the Spokane Education Association found that nearly 25% doubted that principals could conduct fair assessments (2011).

**Lacking evaluator expertise.** A number of studies have revealed concerns about the ability of principals to conduct valid and reliable assessments. According to Donaldson (2009), principals may receive training for the process of evaluation but they rarely have time or opportunity for continuous professional development required to hone their basic skills. Examples of these skills include: (a) ongoing training to calibrate their ratings against other principals’ ratings, (b) discussion with other principals on the meaning of common terms; (c) practice rating a sample observation, and (d) participation in role-playing evaluation conferences.

Himmelein (2009) in a survey of principals’ attitudes toward the evaluation process found that 53% of the administrators who responded to her survey felt they did not receive adequate training to conduct effective teacher evaluations. She also found that 63% of the respondents were employed in districts that did not provide evaluation training for principals. Loup et al. (1996) in a replicative study of 100 large school districts found that 69% of them required teacher evaluation training, but the focus of the training was on procedures and instrumentation and not on improving their personal skills as an evaluator. Brandt, Mathers, Oliva, Brown-Sims, and Hess (2007), in analyzing policies from 140 Midwestern districts found only 11 of them provided written documentation on training requirements for evaluators.

Studies have also revealed inconsistent applications of the evaluation process by individual principals. Little, Goe, and Bell (2009), for example, found that different evaluators rated the same teacher differently, primarily because they had dissimilar views
of effective teaching. Zimmerman and Deckert-Pelton (2003) noted that 38% of the teachers in their study thought the performance evaluations were conducted inconsistently. Donaldson (2012) reported that increased levels of anxiety among teachers who were unwilling to address inconsistency concerns with principals fearing their evaluation ratings would be lowered if they did so.

Institutional restrictions. In a study by Kersten and Israel (2005), 36% of the principals responding cited unions as being impediments to effective teacher evaluations. For example, unions pressured principals to inflate teacher ratings. In a study by Painter (2000), 67% of the responding principals identified the teacher union or the collective bargaining agreement as the greatest barrier to making performance evaluation effective.

In the realm of school culture, Ilgen and Davis (2000) reported that giving teachers negative feedback was often limited because sensitivity to demotivation took precedence over honesty. Halverson, Kelly, and Kimball (2004) in a case study of 14 schools within the same district, reported administrators used teacher performance evaluations to build positive relationships rather than strengthen instructional practices. According to Sawchuck (2013), principals often inflated their ratings of teachers in part because of cultural expectations. He found that high ratings of expected or above expected were given to teachers in the following states even in the context of new rigorous evaluation programs: 98% in Michigan, 94% in Georgia, and 98% in Tennessee. This “leniency bias” was also found to exit in the Cincinnati Public School System with more than 90% of the teachers receiving scores in the two highest categories even after the principals received intensive training to evaluate teacher performance (Taylor & Tyler, 2012).
In summary, teacher dispositions toward performance evaluation may be influenced by a number of common problems including time restrictions, inadequate observations and instrumentation, and the propensity for evaluators to conduct evaluations with bias and subjectivity. These problems contribute to a negative disposition towards performance evaluation, which lessens the impact the evaluation has on the intended goal of improving instruction in order to increase student achievement. Thus, teacher performance evaluation becomes a yearly exercise that yields little benefit to either teachers or their students.

**The Ohio Teacher Evaluation System (OTES)**

**History.** Since 2004, Ohio has made changes in education policy in an effort to improve school performance. Senate Bill 2 in 2004 mandated the creation of the Educator Standards Board, which developed the Ohio Standards for the Teaching Profession, the Ohio Standards for Principals, and the Ohio Standards for Professional Development (NIET, 2012). House Bill 1 in 2009 gave a directive to the Educator Standards Board to recommend model evaluation systems for both principals and teachers to the State Board of Education. In response to this mandate, OTES was officially adopted in House Bill (HB) 153.

As required in HB 153, the following are the system’s defining characteristics:

- Provides for multiple evaluation factors, including student academic growth;
- Is aligned with the standards for teachers adopted under section 3319.61 of the Ohio Revised Code (ORC);
• Requires observation of the teacher being evaluated, including at least two formal observations by the evaluator of at least 30 minutes each and classroom walkthroughs;

• Assigns a rating on each evaluation conducted under sections 3319.111 and 3319.112 of the ORC in accordance with the following levels of performance: Accomplished, Skilled, Developing, or Ineffective;

• Requires each teacher to be provided with a written report of the teacher’s evaluation;

• Implements a classroom-level, value-added program developed by a nonprofit organization as described in division (B) of section 3302.021 of ORC;

• Identifies measures of student academic growth for grade levels and subjects for which the value-added progress dimension prescribed by section 3302.021 of the ORC does not apply;

• Provides for professional development to accelerate and continue teacher growth and provide support to poorly performing teachers; and

• Provides for the allocation of financial resources to support professional development (NIET, 2012, p. 2)

School districts were not required to use the OTES model provided they designed or revised their own evaluation program that complied with HB 153.

After the initial year of implementation in the 2013-14 academic year, legislators responding to teacher and administrative feedback, proposed changes to OTES as outlined in Senate Bill 229. After several months of debate, a compromise resulted in changes as found in HB 362. The changes included extended evaluation cycles for
accomplished and skilled teachers, the choice of an alternate evaluation framework, evaluation exemption for teachers on leave for more than 50% of the school year, and exemption for teachers whose retirement has been approved by the board of education (Ohio House of Representatives, 2014).

**Evaluation components.** OTES in 2013-14 was comprised of two evaluation components: teacher performance rating and student growth measures. The teacher performance rating is derived from scores on the teacher performance rubric based on the *Ohio Standards for the Teaching Profession* and at least two informal walkthroughs during the evaluation cycle. Teachers receive one of the following four ratings:

- **Accomplished rating** indicates that the teacher is a leader and model in the classroom, school, and district exceeding expectations for performance. The teacher consistently strives to improve his or her instructional and professional practice, contributes to the school or district community of practice, and whose students consistently demonstrate above expected academic growth. A teacher receiving accomplished may be placed on a three year evaluation cycle dependent on student progress measures remaining average or above.

- **The skilled rating** indicates that the teacher consistently meets expectations for performance and fully demonstrates all competencies. This rating is the rigorous, expected performance level for most experienced teachers whose students meet or exceed academic student growth expectations. A teacher receiving “skilled” may be placed on a two-year evaluation cycle dependent on student progress measures remaining average or above.
• A rating of developing indicates that the teacher demonstrates minimum competency in many of the teaching standards, but may struggle with others. For example, a teacher who is new to the profession or in a new assignment may be making progress but requires ongoing professional support for necessary growth to occur. An experienced teacher may require more intensive support depending on the areas of improvement. The students of this teacher may meet or exceed expected academic growth but either way there is room for teacher improvement.

• An ineffective rating indicates that the teacher fails to demonstrate minimum competency in one or more teaching standards on a consistent basis. The teacher requires immediate assistance and ongoing support and must be placed on an improvement plan. Students of this teacher demonstrate below expected academic growth. There is little or no improvement over time in performance and/or student academic growth (Ohio Department of Education, 2014).

Student growth measures comprise up to 50% of the teacher performance evaluation (ODE, 2014). Student growth for OTES was determined through a combination of measures: value-added data, vendor created assessments, and Local Education Agency (LEA) measures or student learning objectives (SLOs).

For teachers with value-added data linked to them through state testing results, these results were used for 26-50% of their growth measure (ODE, 2014). The other percentage came from a combination of vendor assessments or SLOs. For teachers without value-added data there were two options: (a) Teachers with approved vendor assessment data used a combination of vendor assessments from an approved ODE vendor list and SLOs; or (b) Teachers without vendor-approved assessment data used
SLO data or a combination of SLO data and shared attribution. Shared attribution, a measure that includes sharing collaborative goals from a grade level, building or district was used from building or district value-added scores, performance index gain, or building or district-based SLOs (NIET, 2012).

New evaluation frameworks. Revisions to OTES were submitted to legislators before the end of the first year of implementation. Beginning with the 2014-15 school year, local districts were allowed to use either the original framework (50% teacher performance and 50% student growth measures) or an alternate framework. The alternate framework is outlined as follows:

- Teacher performance – 42.5% of the summative rating
- Student growth – 42.5% of the summative rating
- One of the following additional measures from an Ohio Department of Education approved list for the remaining 15% of the summative rating:
  - Student surveys
  - Teacher self-evaluation
  - Peer review evaluations
  - Student portfolios

For the 2015-16 school year, the following alternate framework was in effect:

- Teacher performance – 42.5-50% of the summative rating
- Student growth – 42.5-50% of the summative rating
- One of the following approved additional measures for the remaining 1-15%:
  - Student surveys
  - Teacher self-evaluation
Peer review evaluations

Student surveys (ODE, 2014)

Shared responsibilities. The principal and teacher share certain responsibilities in the OTES evaluation process. Conferencing, an important element in the formative evaluation process, is a shared responsibility. During the year, evaluators and teachers have the opportunity for professional conversations about performance, goals, progress, and any needed supports (NIET, 2012). Every formal observation by the evaluator contains both a pre and post conference. The preconference is used to review the formal observation process, discuss the lesson plan, and define specific areas of focus during the observation. At this post conference, the principal and teacher discuss each indicator, determine a course of action for improvement and select areas of reinforcement and refinement.

The final summative rating is a shared responsibility in which evidence is provided by both the evaluator and the teacher from: formal and informal observations, artifacts such as lesson plans and reflection documents, and other evidence collected throughout the cycle (NIET, 2012). Upon examination of the evidence, the evaluator assigns a final rating and then consults with the teacher to build a plan for personal growth or improvement through the discussion of “opportunities for professional development that evolve as a result of the evaluation process” (NIET, 2012, p.12).

Individualized staff development requirements. The Professional Growth Plan, a required element of the OTES process, is intended to help teachers focus on areas of professional development designed to improve their instructional practice (NIET, 2012). At the beginning of the year a professional growth plan or improvement plan is
written based on the previous year’s final summative rating. A growth plan is developed for teachers rated developing, skilled or accomplished. The growth plan includes prior student growth data, feedback from the evaluator, and the teacher’s self-assessment of the seven professional standards (NIET, 2012). Professional development is designed to the individual teacher’s specific area of growth. The evaluator is responsible for providing support and resources for the teacher to act on the professional development opportunities.

Improvement plans are developed for a teacher by an evaluator in response to ratings of ineffectiveness found in performance or student growth. The plan may be initiated at any time during the evaluation cycle based on “deficiencies in performance as documented by evidence collected by the evaluator” (NIET, 2012, p.13). The improvement plan is intended to identify areas of performance improvement while providing guidance and support to the educator.

The responsibility of the administrator for an improvement plan is as follows:

- Identify, in writing, the specific area(s) for improvement to be addressed in relationship to the Ohio Standards for the Teaching Profession;
- Specify, in writing, the desired level of performance that is expected and a reasonable period of time to correct the deficiencies;
- Develop and implement a written plan for improvement that will be initiated immediately and will include resources and assistance available;
- Determine additional education or professional development needed to improve in the identified area(s): and
- Gather evidence of progress or lack of progress (NIET, 2012, p.13)
The teacher’s performance should be reassessed, which is a responsibility of the evaluator. If the improvement has been documented as acceptable, then the regular evaluation cycle will resume. If improvement is not acceptable, the supervising administrator has the option to reinstate the improvement plan with additional recommendations or take necessary steps for a recommendation to dismiss the teacher from their duties.

In summary, OTES contains components that support current trends in effective evaluation practice. These components are a combination of teacher performance, student growth measures, and/or additional measures, i.e. parent and student surveys. Teacher performance and student growth measures are used to determine the development of individualized teacher growth or improvement plans with the goal of improved instruction to positively affect student achievement.

**Using Student Growth Measures to Evaluate Teacher Performance**

Teacher evaluation systems increasingly require the infusion of multiple measures of student growth. These measures seek to determine the level of growth students make over the course of a school term. Research has yet to determine which combination of measures would provide the most accurate and useful information about student growth and its relationship to teacher effectiveness (NIET, 2012).

Many teachers are skeptical or opposed to using student growth measures to evaluate their performance. Sporte et al. (2013) found that 57% of teachers they surveyed stated their evaluation system relied too heavily on student growth measures. Nearly one-third of the 554 teachers in the same study identified the student growth component and assessments as being the most problematic issues in the evaluation
system. Other concerns that were identified include: influences outside the teachers’ control, heavy testing burden, and narrow representation of student learning as measured by standardized tests. Kyriakides and Campbell (1999) interviewed 297 primary teachers and asked their perceptions of incorporating baseline student growth assessments to predict their students’ future progress. Although teachers believe that student data were important, 87% in this study viewed these metrics as the second least important part of the evaluation process.

**Value-added student learning data.** The use of value-added student data is an attempt to address the subjective nature of observation-only teacher performance evaluations. This idea originated in Tennessee in the early 1990s and was viewed skeptically in academic and policy circles for nearly two decades (Saltman, 2010). Peer-reviewed, empirically based research supporting or challenging the implementation of value-added data is growing as the federal Race to the Top grants awarded $4.35 billion to eligible states using this performance measure (Saltman, 2010). Currently 40 states are using value-added models in their teaching performance evaluations (Collins & Amrein-Beardsley, 2014).

Value-added models are used to measure teacher performance or effectiveness with the objective of rewarding or penalizing teachers for the achievement gains their students make (Berliner, 2014). Specifically, value-added metrics measure changes over time in student test scores; changes are then attributed to a teacher’s instruction. If students score higher on a standardized test than they scored in the previous year, the teacher is credited with adding value to student learning. If the test scores do not improve or decline, the opposite is true.
Advocates maintain that value-added models can be combined with other measures, such as observation protocols and student surveys, to assess and promote effective teaching. Value-added models supposedly offer an objective measure of teacher performance that can be quantified and tracked. Value-added assessments attempt to glean from the data those teaching methods which will inform instruction resulting in higher student scores.

Critics claim there are numerous issues with using value-added metrics to determine teacher effectiveness. Concerns relate to the stability of individual teacher value-added scores over time (Goldhaber & Hanson, 2008), nonrandom assignments of students to teachers, and whether the constructs measured by student assessments are consistent or vertically scaled (Rothstein, 2008; Saltman, 2010, Youngs, 2013). Teachers according to Toch and Rothman (2008) are “strongly opposed to evaluations based substantially or exclusively on student test results” (p. 4).

Teachers report it is difficult to consistently maintain high-value added scores. Forty percent or more of the teachers who perform in the top quartile (25%) could not maintain this score the following year (Youngs, 2013). The correlation between an individual teacher’s underlying value-added score based on three or more years of data, and their score for any given year is often much higher that the correlation between value added scores from one year to the next. According to Darling-Hammond (2012) the National Research Council, the Rand Corporation, and Educational Testing Services have recommended that value-added student data should not be used to make high-stakes decisions about teachers because these measures are unreliable, unstable, and biased. According to Berliner (2014), value-added models are not stable enough to be used in
evaluating teachers due to the effects of endless exogenous variables. He asserts identifying the best and worst teachers in a school system using value-added data reveals a morally problematic and psychometrically inadequate base for decision-making (Berliner, 2014).

Konstantopoulos (2014) states value-added models cannot completely eliminate factors that contribute to student achievement such as prior achievement, socioeconomic status, and English language proficiency. Unobserved factors that influence value-added data include parents and the influence of peers. It is difficult to capture the complexity of teaching with this model. Some researchers argue that value-added models should not be used because a large number of teachers do not teach subjects tested on state-mandated assessments (Crouse, 2008). Furthermore, value-added models attribute to a single teacher the teaching done by several teachers.

**Student learning objectives.** SLOs began in the Denver Public School system in 1999 as a means to link student outcomes to teacher pay (Reform Support Network, 2012). Denver’s use of SLOs found that rigorous and high-quality growth objectives were associated with student achievement. SLOs by definition are specific learning goals with a specific measure of student learning used to track progress toward that goal (NIET, 2012). SLOs became a way for states to incorporate student growth measures into their evaluation system for non-tested grades and subjects. Over 72% of state representatives in a survey by Harris and Hofer (2011) reported assessing student progress for teachers without value-added data as found in non-tested grades and subject areas as problematic. Approximately, 30% of teachers fall into this category (Harris & Hofer, 2011).
The literature and research on SLOs is minimal. Nevertheless, research conducted by Gill, Bruch, and Booker (2013) reports the following key findings:

- SLOs have the potential to better distinguish teachers based on performance than traditional evaluation metrics do, but no studies have looked at SLO reliability.
- Little is known about whether SLOs can yield ratings that correlate with other measures of teacher performance. Only three studies have explored the relationship between SLO ratings and standardized assessment-based (value-added) growth measures. These studies found small but positive correlations. More research is needed as states and districts utilize SLOs as teacher evaluation measures and instructional planning tools.
- Until some of the research gaps are filled, districts that intend to use SLOs may want to implement them for instructional planning before using them in high-stakes evaluations. Several studies found teacher concerns about fairness in SLO implementation. This is no surprise because it is difficult to make SLOs that are valid and reliable. They are by definition customized to individual teachers and based on the professional judgments of teachers and principals. Making SLOs an important component of high-stakes evaluation could undermine their validity, because it means that teachers are in essence grading themselves.
- Studies of teacher experiences with SLOs indicate that SLOs can require substantial training and technology infrastructure and that they can be time-consuming for teachers and evaluators alike.
- Principals play a critical role in ensuring consistency and rigor when using SLOs as a measure of student growth. With another area to monitor, this raises
questions of further inconsistency in the entire teacher performance evaluation (p. ii)

Teacher dispositions regarding the use of SLOs were found in a report prepared by the Indiana Department of Education (2012):

- Teachers felt that the SLO process was time-consuming, particularly creating and updating assignments. The assessments should be created or identified prior to the start of the school year.
- Teachers stated it was a challenge to obtain baseline student data from the prior year.
- The majority of teachers believe that SLOs should accompany other measures of student learning in the evaluation system.
- Storing student learning data was problematic. Teachers requested technology solutions to address this problem (pp. 22-24).

In summary, evaluation is the primary medium to affect instructional practice in the classroom. Current evaluation systems promise to deliver improved teacher performance, but fall short due to: negative dispositions and biases towards evaluations as reported by teachers and principals, inadequate evaluation instruments and processes, and the use of questionable student growth measures to rate the effectiveness of instruction. This lack of confidence in teacher performance evaluation by both principals and teachers may stymie professional growth and ultimately student achievement, the goal of teacher performance evaluation.
Linking Performance Evaluation to Individual Growth Plans

The Standards of Professional Learning have been adopted by many states as their guidelines for teacher performance evaluation programs. The standards outline characteristics of professional learning that lead to effective teaching practices, which in turn may lead to improved results (Learning Forward, 2015). A mid-year evaluation is conducted to progress monitor growth. Evidence containing documentation of professional growth activities toward the fulfillment of the growth plans is provided to the evaluator at the end of each cycle (Learning Forward, 2015).

Individual growth plans are different than licensure requirements. For example, districts in Ohio, as mandated by Ohio Revised Code 3319.22, require teachers to have an individual professional development plan (IPDP) that is used to document professional development hours to gain or renew state licensure (ODE, 2014). Individual professional growth plans give new status to professional development through the requirement of individualized professional development as a component of the evaluation process rather than only a means to retain license certification (Allen, 2009).

The growth plans are evidence-based individual growth and development plans organized for teachers to set and pursue professional growth goals that are connected to learning activities as part of their professional review cycle (Minnesota Department of Education, 2013). Depending on the model developed by districts or states, the plan outlines a process that honors differences among teachers while providing support for improving teaching and learning (Chesterfield County Public School System, 2009).

Individualized growth plans reflect best practices in professional development by providing collaboration with peers, discussion of ways to improve teaching and learning
and the opportunity to be reflective as adult learners. According to the National Staff Development Council, sustainability of teacher learning is most likely to occur if the following four conditions exist: 1) learning is in the context of what is taught; 2) it is in conjunction with sustained opportunities to study, experiment with strategies and methods, and receive feedback; 3) teachers have collaboration with professional peers and access to external resources; and 4) teachers have influence over the content and process leading to a sense of ownership (NSDC, 2008). For example, the Chesterfield County Public School plan was developed to “scaffold teachers’ growth by involving them in identifying their professional growth needs and by providing specific useful feedback in their performance” (Chesterfield County Public School System, 2009. p. 3). Glickman, Gordon, and Ross-Gordon (2009) state that adult learners should have targeted professional development that allows for individual teacher goals and self-directed activities. Fenwick (2004) states that plan must meet the needs of the individual teacher in addition to students’ needs.

Teachers who participate in professional development in their content area improve their instructional strategies (Frome, Lasater & Cooney, 2005). Learning that is job-embedded involving group and peer participation is more effective than off-site professional development (Garet et al., 2001). Learning centered, knowledge centered, assessment centered, and community centered activities increase teacher knowledge (Bransford, Brown, & Cocking, 1999). Zepeda (2012) states that professional development should take into account the number of years experience, prior knowledge and career stage for the individual teacher. The individual plan supports these suggestions for adult learning.
Empirical studies in the area of teacher professional growth plans are extremely light due to the newness in this area of study. According to Allen (2009), confidentiality prevents certain aspects of teacher evaluations from being available to researchers. Research conducted on this topic is typically qualitative as evidenced by case studies and interviews.

Teaching experience is a factor in dispositions toward individual professional growth plans. This was evident in the comparison of novice and experienced teachers in a case study by Allen (2009) with the following results:

- Novice teachers reported confusion whereas experienced teachers felt the plan was another form to fill out.
- Teachers stated they were truthful in their documentation of performance results. Conversely, teachers in this same study reported they could name teachers who did not report true results due to the perception that they would receive a low performance evaluation.
- Teachers did not adhere to the requirements of developing their own personal plans for improving instruction. Instead, they based their plans on perceived district and curriculum initiatives for the following year.
- Teachers reported differing opinions on the nature of the relationship between administrators and teachers as a result of personal growth plans. Teachers felt administrators only monitored the plans of teachers who were non-tenured, leaving tenured teachers to self-monitor.

A study by Hughes (2006) concluded that self-evaluation as part of the evaluation plan affected teachers’ professional growth and the relationship between teachers and
their building principals. Teachers aware of their own strengths and weaknesses sought knowledge to improve their teaching skills, which increased contact and improved relationships with their administrators.

Mentors play an important role in professional growth plans. Kahrs and Wells (2012) found growth and development of the teacher required the knowledge and support of unique needs through a mentor. The study focused on beginning teachers with the highest level of self-efficacy as measured on a self-efficacy scale. Key findings showed that while teachers felt support, their dispositions towards growth plans were labeled “confusing” since the role of the mentor had not been clearly defined. Furthermore, lack of specific feedback left the novice teachers craving for meaningful feedback from their mentors.

A qualitative study by Fenwick (2004) supports the need for teacher professional growth plans. Eleven teachers and five administrators from an elementary, junior high, and senior high were interviewed about their experiences with teacher growth plans. The results are as follows:

- All educators claimed that using professional growth plans motivated teachers’ learning. Common beliefs included greater authenticity and teacher commitment to their own professional development, increased teacher focus and accountability for their own development, increased collegiality, and teachers’ self-affirmation.

- All but one teacher felt professional growth plans were self-affirming and claimed to enjoy the principals’ involvement in their learning activities and accomplishments.
• Teachers reported an overall increase in their own learning activity.

• Teachers claimed their involvement in collaborative learning activities (e.g., professional reading, attendance at workshops, joining in-school groups to explore teaching-related topics) had increased since they constructed their first growth plan.

• Teachers stated the value of using professional growth plans increased due to the focus on goal setting.

• Teachers reported that they approached professional development differently with a written, individualized plan.

• Three of the five administrators believed professional growth plans produced greater commitment to ongoing and continuous learning.

• Administrators reported they had a direct influence on goal setting as well as providing targeted resources to meet the goals.

• Four of five administrators noted collaboration increased.

Negative dispositions toward professional growth plans were reported.

• Teachers cited the need for increased time and resources to implement the plan.

• Reflection time was the most frequently named resource teachers desired, but was difficult to obtain.

• Principals saw scheduling time to meet with the teachers as necessary, but difficult to accomplish.

• Teachers described being uncomfortable with a linear planning process applied to professional learning since goals were set at the beginning of
the year. Fenwick (2004) states “Important knowledge related to practice is often emergent and difficult to articulate early in the year” (p. 272).

- Teachers worried about fulfilling their goals. Those that were unable to achieve their goals felt like they had failed.

In a related study, Doherty (2009) studied teachers’ perceptions about the impact of evaluations on professional growth. A negative disposition towards the process was reported. Teachers “did not see the teacher evaluation system as having an effect on professional growth; rather it was other external factors such as professional development offerings, state and federal certification requirements, and contractual salary incentives” (Doherty, 2009, p. 273).

Individual growth plans, when tied to teacher performance evaluation are designed to impact the improvement of instruction. Teachers and administrators value professional growth plans because the plans provide opportunities for goal setting, building relationships between teachers and administrators, motivating learners, and improving practice. Negative dispositions toward teacher performance evaluation were cited in relation to time, anxiety over accomplishing goals, and confusion with the process. Continued research is needed to gain further insight into how professional growth plans as a component of the teacher performance evaluation process may directly improve instruction.

Summary

Student achievement gains have been linked to the quality of instruction in the classroom. The intent of teacher performance evaluation is to identify the quality of instruction in order to improve teacher effectiveness through professional development
opportunities. Evaluations, though, often fall short of identifying effective instructional practice. In Ohio, the Ohio Teacher Evaluation System was created to address concerns with teacher performance evaluation through the use of multiple measures: observation, student growth, and newly developed student and parent surveys. The intent of OTES is to foster teachers’ professional growth through a collaborative and supportive relationship between the teacher and evaluator. OTES is an attempt to bridge the gap between teacher performance evaluation and improved instructional practice.

One barrier to the successful implementation of professional growth may come in the teachers’ dispositions toward OTES. The teacher’s dispositions may affect the evaluation process, relationship with the evaluator, and the level of commitment to improving instructional practices. This study analyzes teachers’ dispositions toward OTES after the first required year of implementation. Data from this proposed study may be used to inform principals in individualizing professional development for teachers, to inform state and school district policy makers and administrators, and contribute to the knowledge base on teacher evaluation and individual growth plans.
CHAPTER III

METHODS

This descriptive, correlational study utilized a survey to collect quantitative data from teachers regarding their dispositions toward OTES. Data were analyzed to determine levels of association between teacher dispositions (the criterion variable) and three predictor variables (gender, years of teaching experience, and grade level teaching assignment) and to determine the extent to which the predictor variables collectively accounted for variability in the criterion variable.

Research Questions

The study was guided by four research questions:

1. What were the respondents’ dispositions toward specific aspects of OTES?
2. What were the individual respondents overall dispositions toward OTES?
3. What was the level of association between the criterion variable (teacher overall personal dispositions toward OTES) and each of three, predictor variables (gender, years of teaching experience, and grade level teaching assignment)?
4. To what extent did the three predictor variables collectively account for the variability in the criterion variable?

Population

Data were obtained from a defined population consisting of teachers (Kindergarten through 12th grade) employed in 8 of the 9 public school districts in Miami
County in Ohio in the 2014-15 school year. One public school district was excluded from the study for non-participation in OTES. The district superintendent cited the teachers’ union contract as the reason the district did not participate.

Miami County was chosen because it contains a representative sample of three types of public school districts found in the state of Ohio: Local District (3), City Districts (2), and Exempted Village Districts (4). Teachers were identified from publicly available Ohio Educational Directory System (OEDS) records. The list included the names and mailing addresses for (a) districts, (b) superintendents, (d) individual schools, and (d) principals. Email addresses for teachers were obtained from school district websites. The total number of teachers in the defined population was 895. The involvement of the entire population was intended to increase the power of the study, which in turn would strengthen the likelihood that the results would be accurate (Heiman, 2011).

**Instrumentation**

The survey instrument was designed by the researcher to collect dispositions toward OTES and respondent demographic information. The survey is an appropriate method for gathering information from a large group of participants using identical procedures (Babbie, 1999; Krathwohl, 2009) when a phenomenon cannot be directly observed. Surveys are used to describe opinions, attitudes, and characteristics (Creswell, 2009). The survey (see Appendix D) took approximately 5 minutes to complete.

The researcher developed the survey instrument using OTES guidelines and studies cited in the literature review (Donaldson, 2012; Halverston, Kelly, & Kimball, 2004; Harris, Rutledge, Ingle, & Thompson, 2010; Ilgen & Davis, 2000; Indiana
The Teachers Dispositions toward OTES Survey consists of 17 Likert-type statements divided into two sections: (1) Elements of OTES and (2) Experiences with OTES. Items were stated both positively and negatively to engage the participant in thoroughly reading each question, which in turn would reduce response bias (Fields, 2005). A five-point scale ranging from Strongly Disagree (1) to Strongly Agree (5) was used to identify level of agreement with each statement. Respondents were provided the choice of Uncertain (3) to describe dispositions on items that did evoke agreement or disagreement. The Likert-type scales used in this study are described as continuous scales, reporting a mean and standard deviation for each statement.

The three remaining survey items pertained to respondent demographic information, specifically gender, teaching experience, and grade level assignment. Responses to these survey items were analyzed using descriptive statistics (e.g., frequencies, ranges, means, and standard deviations as appropriate).

Reliability. Cronbach’s alpha was calculated for the construct of teacher attitudes using scale items to test for internal consistency. The alpha coefficient can range in value from 0 to 1 and was used to describe reliability factors in dichotomous or multi-point scales. The closer the Cronbach alpha score is toward 1, the more reliable the scale. A score of .70 or higher is considered acceptable in most social science research (Santos, 1999).

The sample size utilized in the pilot study was small (\(N = 20\)). Estimates of reliability in a sample size of less than 200 participants are not considered accurate.
Despite the limitations, the reliability estimate was calculated using all scale items to test for internal consistency. All 17 items were used to test for reliability even though the survey was divided into sections, as each section gathered information on the same construct -- teacher dispositions toward OTES. The Cronbach alpha for the instrument when piloted was .70. Following the study, a Cronbach alpha was calculated for the instrument with a demonstrated acceptable reliability level of .70.

Validity. Face and content validity of the instrument were established by a panel of experts as recommended by Krathwohl (2009). The panel assessed the appearance of the instrument and the instrument’s representation of the topic (Burton & Mazerolle, 2011). The panel members are listed in Appendix A. The survey was also reviewed by the author’s dissertation committee. The survey was revised based on the recommendations of the panel and dissertation committee.

Construct validity for teacher dispositions was calculated using several methods. OTES, which is based on the Ohio Standards for Teachers, was used to create items for this survey. Extensive research had been conducted on the Ohio Standards for Teachers in order to set performance expectations. Survey questions were developed that reflect the elements of OTES (See Table 1).
Table 1

Survey Instrument Questions Developed from OTES Elements

<table>
<thead>
<tr>
<th>OTES Elements</th>
<th>Question #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher performance evaluation</td>
<td>7, 9</td>
</tr>
<tr>
<td>Walkthroughs</td>
<td>8</td>
</tr>
<tr>
<td>Individual Growth Improvement Plan</td>
<td>1, 2</td>
</tr>
<tr>
<td>Student Growth Measures</td>
<td>3, 4, 5, 6</td>
</tr>
<tr>
<td>Value-added</td>
<td>3</td>
</tr>
<tr>
<td>SLOs</td>
<td>4</td>
</tr>
<tr>
<td>Vendor Assessments</td>
<td>5</td>
</tr>
<tr>
<td>Alternate Assessments</td>
<td>6</td>
</tr>
<tr>
<td>(Student survey, peer-evaluation, teacher reflection, student portfolios)</td>
<td></td>
</tr>
</tbody>
</table>

Validity was further established through several studies (Himmelein, 2009; Loup et. al, 1996; Zimmerman & Deckert-Pelton, 2003), which used surveys or interview questions specifically designed to gather data on teacher dispositions as a guide in the development of the survey questions (Table 2). Evidence of the ability of the instrument to measure dispositions was obtained by piloting the survey instrument and establishing construct validity prior to administration as recommended by Krathwohl (2009). The pilot consisted of 20 teachers outside the study population.
Table 2

*Research Source of Survey Instrument Questions*

<table>
<thead>
<tr>
<th>Source</th>
<th>Question #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calabrese, Sherwood, Fast &amp; Womack, 2004</td>
<td>17</td>
</tr>
<tr>
<td>Donaldson, 2013</td>
<td>16</td>
</tr>
<tr>
<td>Ilgen &amp; Davis, 2000; Halverston, Kelly &amp; Kimball, 2004</td>
<td>10</td>
</tr>
<tr>
<td>OTES, 2013</td>
<td>1, 2, 3, 4, 5, 6, 8, 12, 14</td>
</tr>
<tr>
<td>OTES, 2014</td>
<td>7, 9, 11</td>
</tr>
<tr>
<td>Sporte, et al., 2014</td>
<td>13</td>
</tr>
<tr>
<td>Toch &amp; Rothman, 2008</td>
<td>15</td>
</tr>
</tbody>
</table>

**Data Collection**

The survey was conducted using Survey Monkey, an electronic survey instrument. All participants were requested to participate via an email sent to their respective school building addresses. The initial email described the purpose of the research, requested individual teacher participation, and ensured anonymity (see Appendix C). Each individual who accepted the invitation was sent (via email) the survey instrument (see Appendix D).

Data were collected in May and June, 2015. This time frame was deemed advantageous because principals had to complete OTES by May 1, 2015. Members of the study population were provided a statement explaining that completion and return of the survey implies consent to participate in the study (see Appendix C).
instruments were returned via email to the researcher. Participants were not compensated for their participation.

Within 14 days after receiving the survey, 85 teachers (9%) responded. A reminder email was then sent to non-responders. The reminder resulted in an additional 57 completed surveys, or a total of 142 respondents. No surveys were completed after June 14, 2015.

**Data Analysis**

Data analysis was conducted using descriptive statistics (mean, standard deviation, frequency counts, and percentages) to document responses to statement or question on the survey. Additional analysis of the data to potentially clarify relationships was then conducted. Each research question is listed with the associated data analysis explained.

Research question one: *What were the respondents’ dispositions toward specific aspects of OTES?* To answer this question, responses to the 17 statements were assigned numeric values. Responses to the positively-keyed statements (2, 3, 8, 9, 10, 11, 12, 14, 15, and 16) were assigned the following values:

- *Strongly disagree* 1
- *Disagree* 2
- *Uncertain* 3
- *Agree* 4
- *Strongly Agree* 5

Negatively-keyed statements (1, 4, 5, 6, 7, 13, & 17) were reverse coded and thus, they were assigned the following values:
Strongly disagree 5
Disagree 4
Uncertain 3
Agree 2
Strongly Agree 1

Higher scores, for individual statements reflect a positive disposition toward OTES. This scoring ensures the responses are consistent with other in terms of what the assigned values imply (Psych.wfu.edu., 2013). A mean score was determined for each statement. Statements were categorized based on mean scores using a 5-category typology:

1. **Positive disposition**, a mean score between 4.0 and 5.0
2. **Moderately positive disposition**, a mean score between 3.01 and 3.99
3. **Neither positive nor negative**, a mean score of 3.0
4. **Moderately negative disposition**, a mean score of 2.0 and 2.99
5. **Negative disposition**, a mean score between 1.0 and 1.99.

Other descriptive statistics (frequency counts, percentages, cumulative percentages and standard deviations) were calculated for responses to each statement.

Research question two: *What were the individual respondents’ overall dispositions toward OTES?* To answer this question, a total score for each respondent’s overall disposition was calculated. First, the responses to all 17 items were added to obtain a total score for each respondent. The potential range of total scores for a respondent was 17 to 85. A four-category typology was then created. Respondent’s scores were placed in one of the following categories:
1. **Negative disposition** 17 to 34 points
2. **Moderately Negative disposition** 22 to 43 points
3. **Moderately Positive disposition** 44 to 64 points
4. **Positive disposition** 65 to 85 points

Research question three: *What is the level of association between the criterion variable (teacher overall personal dispositions toward OTES) and each of three predictor variables (gender, years of teaching experience, and grade level teaching assignment)?* To answer this question, gender was recorded as female or male. Teaching experience was recorded as the actual number of years for each respondent.

Grade level teaching assignments were categorized as follows:

- Primary Teacher (Grades K-2)
- Intermediate Teacher (Grades 3-5)
- Middle School Teacher (Grades 6, 7, & 8)
- High School Teacher (Grades 9, 10, 11, & 12)

The *criterion* variable, teachers’ overall personal dispositions toward OTES, was calculated by assigning point values to each response and totaling the points for the 17 statements.

A Pearson product moment correlation, symbolized by the letter “r,” was calculated to measure the strength and direction of the relationship between teacher overall personal dispositions toward OTES (the criterion variable) and each of three predictor variables. One predictor variable, gender, is a nominal, binary variable. An appropriate method for calculating the coefficient for gender and dispositions (an interval variable) is the point-biserial correlation coefficient. A point-biserial correlation ($r_{bp}$) is
used to estimate the degree of relationship between a naturally occurring binary variable and an interval or ratio scale (Cohen, Cohen, West, & Aiken, 2003). SPSS does not have a special procedure for point-biserial correlations, therefore, the use of the Pearson’s $r$ is considered appropriate (Statistics Solutions, n.d.)

The correlation coefficient can range in value from +1 to -1. An increase in the value of one variable accompanied by the increase in the value in the other variable indicates a positive relationship. If the value in one variable increases and the other decreases, the relationship is negative. The researcher utilized the following rubric described by Pallant (2007) to determine the strength of association:

- Small association: (+ or -) correlations from .01 to .29
- Moderate association: (+ or -) correlations from .30 to .49
- Large association: (+ or -) correlations from .50 and higher

The estimates of effect size (association) are relative only to each other and for this reason, findings of association should not be generalized beyond the study population.

Research question four: To what extent did the three predictor variables collectively account for the variability in the criterion variable? A multiple correlation coefficient (R) was calculated to determine the association between the criterion variable (teacher overall personal dispositions toward OTES) and the predictor variables (gender, years of teaching experience, and grade level assignment) collectively. The coefficient was then squared to produce the coefficient of determination ($R^2$), a statistic that determines the proportion of common variation attributable to the predictor variables collectively. The $R^2$ was adjusted to account for the three predictor variables.
Protection of Human Subjects

This research project was reviewed and approved by the University of Dayton’s Institutional Review Board on April 9, 2015. All appropriate ethical standards and statutory requirements were applied in the collection of data from study participants and in the execution of this research.
CHAPTER IV
REPORT OF FINDINGS

This chapter begins with a demographic profile of the respondents. The remainder of the chapter is organized chronologically based on the four research questions presented in Chapter 1. The purposes of this chapter are to report findings pertaining to: (a) respondent dispositions toward specific aspects of OTES, (b) individual respondents’ overall dispositions toward OTES, (c) levels of association between the criterion variable (teacher overall personal dispositions toward OTES) and each of three predictor variables (gender, years of teaching experience, and level of teaching assignment), and (d) the extent to which the predictor variables collectively accounted for variability in teachers’ dispositions.

Respondents’ Profile

Findings presented in this chapter are the result of data collected and analyzed from a defined population consisting of all Miami County public school teachers who were evaluated under OTES during the 2014-15 school year (N = 895). Data were collected utilizing the Teachers’ Dispositions toward OTES Survey, an instrument designed by the researcher for this study. Data collection occurred over a one-month period between May 14, 2015 and June 14, 2015. One hundred forty-two completed surveys were returned and analyzed, a return rate of 15.8%.

The profile, based on the three predictor variables, is in Table 3
Table 3

Respondents’ Profile (N = 142)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Response</th>
<th>Frequency</th>
<th>Range</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Options</td>
<td>(%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>112</td>
<td>(78.9%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>29</td>
<td>(20.4%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teaching</td>
<td>Experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years</td>
<td>142</td>
<td>(100%)</td>
<td>1 to 37</td>
<td>14.3</td>
<td>8.48</td>
</tr>
<tr>
<td>Grade Level</td>
<td>Assignment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K-2</td>
<td>29</td>
<td>(20.4%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-5</td>
<td>32</td>
<td>(22.5%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-8</td>
<td>29</td>
<td>(20.4%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9-12</td>
<td>52</td>
<td>(36.6%)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Findings Related to Research Questions

**Research question 1.** *What were the respondents’ dispositions toward specific aspects of OTES?* To answer the first research question, teachers’ dispositions toward OTES were determined. Respondents identified their level of agreement with statements comprising survey items 1-17 on the *Teachers’ Dispositions toward OTES Survey*. Five response choices on a Likert type scale from which the respondents selected one choice: *(strongly disagree, disagree, uncertain, agree, strongly agree)*, with a point value of 1 assigned to strongly disagree, 2 assigned to disagree, 3 assigned to uncertain, 4 assigned to agree, and 5 assigned to strongly agree.

As described in the previous chapter, some statements were positively-keyed and others were negatively-keyed to reduce response bias. By assigning numeric values to responses, mean scores were calculated for each of the 17 statements; the range of possible means was 1.0 to 5.0. Dispositions toward each statement were interpreted according to the following values:

- *Positive disposition*, a mean score between 4.0 and 5.0
- *Moderately positive disposition*, a mean score between 3.01 and 3.99
- *Neither positive nor negative*, a mean score of 3.0
- *Moderately negative disposition*, a mean score of 2.0 and 2.99
- *Negative disposition*, a mean score between 1.0 and 1.99.

Response data for the 17 specific aspects are in Tables 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, and 20.
Table 4

*Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142)*

<table>
<thead>
<tr>
<th>Statement</th>
<th>(Coding Value)</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>(5)</td>
<td>11</td>
<td>7.7</td>
<td>7.0</td>
</tr>
<tr>
<td>Disagree</td>
<td>(4)</td>
<td>88</td>
<td>62.0</td>
<td>69.7</td>
</tr>
<tr>
<td>Uncertain</td>
<td>(3)</td>
<td>16</td>
<td>11.3</td>
<td>81.0</td>
</tr>
<tr>
<td>Agree</td>
<td>(2)</td>
<td>16</td>
<td>11.3</td>
<td>92.3</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>(1)</td>
<td>11</td>
<td>7.7</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>142</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

This statement is negatively-keyed. The mean response score for the statement is 3.50 ($M = 3.50$) indicating a moderately positive disposition. The standard deviation is 1.05 ($SD = 1.05$).
Table 5

*Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142)*

<table>
<thead>
<tr>
<th>Statement</th>
<th>Coding Value</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>(1)</td>
<td>4</td>
<td>2.8</td>
<td>2.8</td>
</tr>
<tr>
<td>Disagree</td>
<td>(2)</td>
<td>21</td>
<td>14.8</td>
<td>17.6</td>
</tr>
<tr>
<td>Uncertain</td>
<td>(3)</td>
<td>29</td>
<td>20.4</td>
<td>38.0</td>
</tr>
<tr>
<td>Agree</td>
<td>(4)</td>
<td>59</td>
<td>41.5</td>
<td>79.6</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>(5)</td>
<td>29</td>
<td>20.5</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>142</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

This statement is positively-keyed. The mean response score for the statement is 3.61 ($M = 3.61$) indicating a moderately positive disposition. The standard deviation is 1.05 ($SD = 1.05$).
Table 6

*Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142)*

<table>
<thead>
<tr>
<th>Statement</th>
<th>(Coding Value)</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>(1)</td>
<td>79</td>
<td>55.6</td>
<td>55.6</td>
</tr>
<tr>
<td>Disagree</td>
<td>(2)</td>
<td>40</td>
<td>28.2</td>
<td>83.8</td>
</tr>
<tr>
<td>Uncertain</td>
<td>(3)</td>
<td>14</td>
<td>9.9</td>
<td>93.7</td>
</tr>
<tr>
<td>Agree</td>
<td>(4)</td>
<td>8</td>
<td>5.6</td>
<td>99.3</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>(5)</td>
<td>1</td>
<td>.7</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>142</td>
<td>100</td>
<td>100.0</td>
</tr>
</tbody>
</table>

This statement is positively-keyed. The mean response score for the statement is 1.67 ($M = 1.67$) indicating a negative disposition. The standard deviation is .91 ($SD = .91$).
Table 7

*Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142)*

<table>
<thead>
<tr>
<th>Statement</th>
<th>Coding Value</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>(5)</td>
<td>9</td>
<td>6.3</td>
<td>6.3</td>
</tr>
<tr>
<td>Disagree</td>
<td>(4)</td>
<td>28</td>
<td>19.7</td>
<td>26.0</td>
</tr>
<tr>
<td>Uncertain</td>
<td>(3)</td>
<td>22</td>
<td>15.5</td>
<td>41.5</td>
</tr>
<tr>
<td>Agree</td>
<td>(2)</td>
<td>37</td>
<td>26.1</td>
<td>67.6</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>(1)</td>
<td>46</td>
<td>32.4</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>142</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

This statement is negatively-keyed. The mean response score for the statement is 2.41 ($M = 2.41$) indicating a moderately negative disposition. The standard deviation is 1.29 ($SD = 1.29$).
Table 8

*Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142)*

<table>
<thead>
<tr>
<th>Statement</th>
<th>(Coding Value)</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>(5)</td>
<td>2</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td>Disagree</td>
<td>(4)</td>
<td>5</td>
<td>3.5</td>
<td>4.9</td>
</tr>
<tr>
<td>Uncertain</td>
<td>(3)</td>
<td>17</td>
<td>12.0</td>
<td>16.9</td>
</tr>
<tr>
<td>Agree</td>
<td>(2)</td>
<td>42</td>
<td>29.6</td>
<td>46.5</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>(1)</td>
<td>76</td>
<td>53.5</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>142</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

This statement is negatively-keyed. The mean response score for the statement is 1.69 ($M = 1.69$) indicating a negative disposition. The standard deviation is .91 ($SD = .91$).
Table 9

*Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142)*

Statement 6. The approved alternate framework (42.5% teacher performance; 42.5% student growth; and 15% additional measure) should not be used to evaluate a teacher’s performance.

<table>
<thead>
<tr>
<th>(Coding Value)</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>(5)</td>
<td>3</td>
<td>2.1</td>
</tr>
<tr>
<td>Disagree</td>
<td>(4)</td>
<td>5</td>
<td>3.5</td>
</tr>
<tr>
<td>Uncertain</td>
<td>(3)</td>
<td>20</td>
<td>14.1</td>
</tr>
<tr>
<td>Agree</td>
<td>(2)</td>
<td>52</td>
<td>36.6</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>(1)</td>
<td>62</td>
<td>43.7</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>142</td>
<td>100.0</td>
</tr>
</tbody>
</table>

This statement is negatively-keyed. The mean response score for the statement is 1.83 ($M = 1.83$) indicating a negative disposition. The standard deviation is .94 ($SD = .94$).
Table 10

*Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142)*

Statement 7. The required level of conferencing between a teacher and the principal (or other evaluator) is excessive.

<table>
<thead>
<tr>
<th>(Coding Value)</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>(5)</td>
<td>4</td>
<td>2.8</td>
</tr>
<tr>
<td>Disagree</td>
<td>(4)</td>
<td>69</td>
<td>48.6</td>
</tr>
<tr>
<td>Uncertain</td>
<td>(3)</td>
<td>11</td>
<td>7.7</td>
</tr>
<tr>
<td>Agree</td>
<td>(2)</td>
<td>40</td>
<td>28.2</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>(1)</td>
<td>18</td>
<td>12.7</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>142</td>
<td>100.0</td>
</tr>
</tbody>
</table>

This statement is negatively-keyed. The mean response score for the statement is 3.00 ($M$ = 3.00) indicating neither a positive nor negative disposition. The standard deviation is 1.18 ($SD$ = 1.18).
Table 11

*Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142)*

Statement 8. Walk-through observations conducted by the principal (or other evaluator) should be used to evaluate teacher performance.

<table>
<thead>
<tr>
<th>Coding Value</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>(1)</td>
<td>12</td>
<td>8.4</td>
</tr>
<tr>
<td>Disagree</td>
<td>(2)</td>
<td>12</td>
<td>8.4</td>
</tr>
<tr>
<td>Uncertain</td>
<td>(3)</td>
<td>17</td>
<td>12.0</td>
</tr>
<tr>
<td>Agree</td>
<td>(4)</td>
<td>89</td>
<td>62.8</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>(5)</td>
<td>12</td>
<td>8.4</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>142</td>
<td>100.0</td>
</tr>
</tbody>
</table>

This statement is positively-keyed. The mean response score for the statement is 3.54 ($M = 3.54$) indicating a positive disposition. The standard deviation is 1.04 ($SD = 1.04$).
Table 12  

*Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142)*

<table>
<thead>
<tr>
<th>Statement</th>
<th>(Coding Value)</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>(1)</td>
<td>20</td>
<td>14.0</td>
<td>14.0</td>
</tr>
<tr>
<td>Disagree</td>
<td>(2)</td>
<td>61</td>
<td>43.0</td>
<td>57.0</td>
</tr>
<tr>
<td>Uncertain</td>
<td>(3)</td>
<td>18</td>
<td>12.7</td>
<td>69.7</td>
</tr>
<tr>
<td>Agree</td>
<td>(4)</td>
<td>42</td>
<td>29.6</td>
<td>99.3</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>(5)</td>
<td>1</td>
<td>.7</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>142</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

This statement is positively-keyed. The mean response score for the statement is 2.59 ($M = 2.59$) indicating a moderately negative disposition. The standard deviation is 1.07 ($SD = 1.07$).
Table 13

*Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142)*

<table>
<thead>
<tr>
<th>Statement</th>
<th>(Coding Value)</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree (1)</td>
<td>29</td>
<td>20.4</td>
<td>20.4</td>
<td></td>
</tr>
<tr>
<td>Disagree (2)</td>
<td>51</td>
<td>35.9</td>
<td>56.3</td>
<td></td>
</tr>
<tr>
<td>Uncertain (3)</td>
<td>42</td>
<td>29.6</td>
<td>85.9</td>
<td></td>
</tr>
<tr>
<td>Agree (4)</td>
<td>18</td>
<td>12.7</td>
<td>98.6</td>
<td></td>
</tr>
<tr>
<td>Strongly Agree (5)</td>
<td>2</td>
<td>1.4</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>142</td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This statement is positively-keyed. The mean response score for the statement is $2.38 (M = 2.38)$ indicating a moderately negative disposition. The standard deviation is $.99 (SD = .99).
Table 14

*Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142)*

<table>
<thead>
<tr>
<th>Statement</th>
<th>Coding Value</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strongly Disagree</strong></td>
<td>(1)</td>
<td>9</td>
<td>6.3</td>
<td>6.3</td>
</tr>
<tr>
<td><strong>Disagree</strong></td>
<td>(2)</td>
<td>15</td>
<td>10.6</td>
<td>16.9</td>
</tr>
<tr>
<td><strong>Uncertain</strong></td>
<td>(3)</td>
<td>24</td>
<td>16.9</td>
<td>33.8</td>
</tr>
<tr>
<td><strong>Agree</strong></td>
<td>(4)</td>
<td>73</td>
<td>51.4</td>
<td>85.2</td>
</tr>
<tr>
<td><strong>Strongly Agree</strong></td>
<td>(5)</td>
<td>21</td>
<td>14.8</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>142</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

This statement is positively-keyed. The mean response score for the statement is 3.57 ($M = 3.57$) indicating a moderately positive disposition. The standard deviation is 1.06 ($SD = 1.06$).
Table 15

*Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142)*

<table>
<thead>
<tr>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>OTES increased the accuracy of evaluations.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(Coding Value)</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>36</td>
<td>25.4</td>
<td>25.4</td>
</tr>
<tr>
<td>Disagree</td>
<td>62</td>
<td>43.7</td>
<td>69.1</td>
</tr>
<tr>
<td>Uncertain</td>
<td>37</td>
<td>26.0</td>
<td>95.1</td>
</tr>
<tr>
<td>Agree</td>
<td>6</td>
<td>4.2</td>
<td>99.3</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>1</td>
<td>.7</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>142</td>
<td></td>
<td>100.0</td>
</tr>
</tbody>
</table>

This statement is positively-keyed. The mean response score for the statements is 2.11 ($M = 2.11$) indicating a moderately negative disposition. The standard deviation is .85 ($SD = .85$).
Table 16

*Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142)*

<table>
<thead>
<tr>
<th>Statement</th>
<th>(Coding Value)</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>(5)</td>
<td>6</td>
<td>4.2</td>
<td>4.2</td>
</tr>
<tr>
<td>Disagree</td>
<td>(4)</td>
<td>41</td>
<td>28.9</td>
<td>33.1</td>
</tr>
<tr>
<td>Uncertain</td>
<td>(3)</td>
<td>32</td>
<td>22.5</td>
<td>55.6</td>
</tr>
<tr>
<td>Agree</td>
<td>(2)</td>
<td>46</td>
<td>32.4</td>
<td>88</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>(1)</td>
<td>17</td>
<td>12.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>142</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

This statement is negatively-keyed. The mean response score for the statement is 2.80 ($M = 2.80$) indicating a moderately negative disposition. The standard deviation is 1.11 ($SD = 1.11$).
Table 17

*Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142)*

<table>
<thead>
<tr>
<th>Statement</th>
<th>(Coding Value)</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>(1)</td>
<td>47</td>
<td>33.1</td>
<td>33.3</td>
</tr>
<tr>
<td>Disagree</td>
<td>(2)</td>
<td>52</td>
<td>36.6</td>
<td>69.7</td>
</tr>
<tr>
<td>Uncertain</td>
<td>(3)</td>
<td>24</td>
<td>16.9</td>
<td>86.6</td>
</tr>
<tr>
<td>Agree</td>
<td>(4)</td>
<td>18</td>
<td>12.7</td>
<td>99.3</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>(5)</td>
<td>1</td>
<td>.7</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>142</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

This statement is positively-keyed. The mean response score for the statement is 2.11 ($M = 2.78$) indicating a moderately negative disposition. The standard deviation is 1.03 ($SD = 1.03$).
Table 18

*Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142)*

Statement 15. The principal (or other evaluator) assisted teachers to develop individual growth plans.

<table>
<thead>
<tr>
<th>Coding Value</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>(1)</td>
<td>25</td>
<td>17.6</td>
</tr>
<tr>
<td>Disagree</td>
<td>(2)</td>
<td>38</td>
<td>26.8</td>
</tr>
<tr>
<td>Uncertain</td>
<td>(3)</td>
<td>26</td>
<td>18.3</td>
</tr>
<tr>
<td>Agree</td>
<td>(4)</td>
<td>49</td>
<td>34.5</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>(5)</td>
<td>4</td>
<td>2.8</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>142</td>
<td>100.0</td>
</tr>
</tbody>
</table>

This statement is positively-keyed. The mean response score for the statement is 2.78 (\(M = 2.78\)) indicating a moderately negative disposition. The standard deviation is 1.17 (\(SD = 1.17\)).
Table 19

*Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142)*

<table>
<thead>
<tr>
<th>Statement</th>
<th>(Coding Value)</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>(1)</td>
<td>24</td>
<td>16.9</td>
<td>16.9</td>
</tr>
<tr>
<td>Disagree</td>
<td>(2)</td>
<td>46</td>
<td>32.4</td>
<td>49.3</td>
</tr>
<tr>
<td>Uncertain</td>
<td>(3)</td>
<td>49</td>
<td>34.5</td>
<td>83.8</td>
</tr>
<tr>
<td>Agree</td>
<td>(4)</td>
<td>22</td>
<td>15.5</td>
<td>99.3</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>(5)</td>
<td>1</td>
<td>.7</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>142</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

This statement is positively-keyed. The mean response score for the statement is 2.50 ($M = 2.50$) indicating a moderately negative disposition. The standard deviation is .97 ($SD = .97$).
Table 20

*Frequency, Percent, and Cumulative Percent for Teachers’ Dispositions (N = 142)*

<table>
<thead>
<tr>
<th>Statement</th>
<th>(Coding Value)</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>(5)</td>
<td>3</td>
<td>2.1</td>
<td>2.1</td>
</tr>
<tr>
<td>Disagree</td>
<td>(4)</td>
<td>6</td>
<td>4.2</td>
<td>6.3</td>
</tr>
<tr>
<td>Uncertain</td>
<td>(3)</td>
<td>7</td>
<td>4.9</td>
<td>11.2</td>
</tr>
<tr>
<td>Agree</td>
<td>(2)</td>
<td>48</td>
<td>34.8</td>
<td>46.0</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>(1)</td>
<td>78</td>
<td>54.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>142</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

This statement is negatively-keyed. The mean response score for statement is 1.64 (\(M = 1.64\)) indicating a negative disposition. The standard deviation is .91 (\(SD = .91\)).

Overall, respondents’ dispositions toward the 17 aspects of OTES addressed in this study were more negative than positive. Specifically, only four dispositions were *moderately positive*, one was neither *positive* nor *negative*, 12 were *negative*, and none of the dispositions were *positive*. The mean and classification for all 17 aspects are included in Table 21 and listed in descending order.
Table 21

*Means and Disposition Classifications for Seventeen Specific Aspects of OTES in Descending Order by Means (N = 142)*

<table>
<thead>
<tr>
<th>Statement</th>
<th>M</th>
<th>Disposition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual staff development is more effective than school-wide development</td>
<td>3.61</td>
<td>Moderately positive</td>
</tr>
<tr>
<td>The principal (or other evaluator) followed the OTES instructions correctly</td>
<td>3.57</td>
<td>Moderately positive</td>
</tr>
<tr>
<td>Walk-through observations conducted by the principal (or other evaluator) should be used to evaluate teacher performance</td>
<td>3.54</td>
<td>Moderately positive</td>
</tr>
<tr>
<td>Requiring teachers to develop annual growth plans is unreasonable</td>
<td>3.50</td>
<td>Moderately positive</td>
</tr>
<tr>
<td>The required level of conferencing between a teacher and the principal (or other evaluator) is excessive</td>
<td>3.00</td>
<td>Neither positive or negative</td>
</tr>
<tr>
<td>OTES had a negative effect on principal (or other evaluator)-teacher relationships</td>
<td>2.80</td>
<td>Moderately negative</td>
</tr>
<tr>
<td>The principal (or other evaluator) assisted teachers to develop individual growth plans</td>
<td>2.78</td>
<td>Moderately negative</td>
</tr>
<tr>
<td>Teacher's instructions for applying OTES were clear</td>
<td>2.59</td>
<td>Moderately negative</td>
</tr>
<tr>
<td>The level of evaluator subjectivity under OTES</td>
<td>2.50</td>
<td>Moderately negative</td>
</tr>
</tbody>
</table>
was low.
Compared to previous evaluation methods, evaluator bias was less common because of OTES.
Student learning objectives (SLOs) should not be used to evaluate a teacher’s performance.
Staff development was more individualized because of OTES.
OTES increased the accuracy of evaluations.
The approved alternate framework (42.5% teacher performance; 42.5% student growth; and 15% additional measure) should not be used to evaluate a teacher’s performance.
Vendor assessment scores should not be used to evaluate a teacher’s performance.
Student value-added measures should be used to evaluate a teacher’s performance.
The amount of time teachers had to devote to OTES was excessive.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Score</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compared to previous evaluation methods, evaluator bias was less common</td>
<td>2.38</td>
<td>Moderately negative</td>
</tr>
<tr>
<td>because of OTES.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student learning objectives (SLOs) should not be used to evaluate a</td>
<td>2.14</td>
<td>Moderately negative</td>
</tr>
<tr>
<td>teacher’s performance.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staff development was more individualized because of OTES.</td>
<td>2.11</td>
<td>Moderately negative</td>
</tr>
<tr>
<td>OTES increased the accuracy of evaluations.</td>
<td>2.11</td>
<td>Moderately negative</td>
</tr>
<tr>
<td>The approved alternate framework (42.5% teacher performance; 42.5%</td>
<td>1.83</td>
<td>Negative</td>
</tr>
<tr>
<td>student growth; and 15% additional measure) should not be used to</td>
<td></td>
<td></td>
</tr>
<tr>
<td>evaluate a teacher’s performance.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vendor assessment scores should not be used to evaluate a teacher’s</td>
<td>1.69</td>
<td>Negative</td>
</tr>
<tr>
<td>performance.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student value-added measures should be used to evaluate a teacher’s</td>
<td>1.67</td>
<td>Negative</td>
</tr>
<tr>
<td>performance.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The amount of time teachers had to devote to OTES was excessive.</td>
<td>1.64</td>
<td>Negative</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Research question 2.** What were the individual respondents’ overall dispositions toward OTES? To answer this question, an individual respondent’s overall disposition toward OTES was determined by assigning point values to each response for the 17 statements as specified in the previous section addressing the first research question and then adding the points. The range of possible individual respondent scores was 17 to 85 and categorization was based on the following rubric:

- *Negative disposition* 17 to 34 points
- *Moderately negative disposition* 35 to 51 points
- *Moderately positive disposition* 52 to 68 points
- *Positive disposition* 69 to 85 points

Data in Table 22 show a majority of respondents (85.9%) had a negative or moderately negative overall disposition toward OTES, with no respondents exhibiting a positive disposition toward OTES.

Table 22

*Overall Respondents’ Dispositions by Categories (N = 142)*

<table>
<thead>
<tr>
<th>Overall Disposition</th>
<th>N</th>
<th>%</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>17-34  Negative</td>
<td>15</td>
<td>10.6</td>
<td>10.6</td>
</tr>
<tr>
<td>35-51 Moderately Negative</td>
<td>197</td>
<td>75.3</td>
<td>85.9</td>
</tr>
<tr>
<td>52-68 Moderately Positive</td>
<td>20</td>
<td>14.1</td>
<td>100.0</td>
</tr>
<tr>
<td>69-85 Positive</td>
<td>0</td>
<td>0.0</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>142</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>
Research question 3. What was the level of association between the criterion variable (teacher overall personal dispositions toward OTES) and each of three, predictor variables (gender, years of teaching experience, and grade level teaching assignment)?

To answer this question, a Pearson product-moment correlation coefficient was calculated using the criterion variable (teacher overall personal dispositions toward OTES) and data for each demographic variable provided by respondents on the survey. The researcher utilized the following rubric described by Pallant (2007) to classify strength of association:

- Small association: (+ or -) correlations from .01 to .29
- Moderate association: (+ or -) correlations from .30 to .49
- Large association: (+ or -) correlations of .50 or higher.

The level of association in all three cases was small and negative. When the \( r \) is close to 0, there is little association between the criterion variable and the predictor variables. A negative association indicates that as the value of one variable increases, the value of the other variable decreases. Therefore, little or no association existed between teachers’ dispositions and any of the three predictor variables. The findings are presented in Table 23.
Table 23

Correlation Coefficient and Levels of Association between Teachers’ Dispositions and the Three Predictor Variables

<table>
<thead>
<tr>
<th></th>
<th>Correlation Coefficient</th>
<th>Descriptive Level of Association</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>-.01</td>
<td>Small, negative</td>
</tr>
<tr>
<td>Years of Teaching Experience</td>
<td>-.16</td>
<td>Small, negative</td>
</tr>
<tr>
<td>Grade Level Assignment</td>
<td>-.19</td>
<td>Small, negative</td>
</tr>
</tbody>
</table>

**Research question 4.** To what extent did the three, predictor variables collectively account for the variability in the criterion variable? A multiple correlation coefficient ($R$) was calculated to determine the association between the criterion variable (teacher overall personal dispositions toward OTES) and the three predictor variables (gender, years of teaching experience, and grade level assignment) collectively. The multiple, correlation coefficient ($R$) was .256. The coefficient was squared to establish the coefficient of determination ($R^2$) a statistic identifying the proportion of common variance produced by the predictor variables; this value was .045.

The ($R^2$) was adjusted to account for the small sample size. According to Pallant (2010), an $R^2$ value tends to be an over-estimation of value when a small sample is involved. As a result an adjusted $R^2$ corrects this value and provides a better estimation of the true population value. In summary, the multiple, correlation coefficient ($R$) was .256, the value of $R^2$ was .065, and the adjusted $R^2$ was .045. Accordingly, the three
demographic variables accounted for only 4.5% of the variability in teachers’ dispositions.

**Summary**

Findings are based on 142 completed and analyzed surveys. A respondents’ profile was created utilizing gender, years of teaching experience and grade level assignment. Data were analyzed to determine teachers’ dispositions toward each of the 17 specific aspects of OTES and the overall disposition of each respondent towards OTES. Level of associations between **teacher overall personal dispositions toward OTES** (criterion variable) and each of three predictor variables (**gender, years of teaching experience, and grade level teaching assignment**) were found to be small and negative. Collectively, the three, predictor variables accounted for only 4.5% of variance in the teachers’ overall dispositions toward OTES.
CHAPTER V
SUMMARY OF FINDINGS, CONCLUSIONS AND RECOMMENDATIONS

The main purpose of this study was to determine teachers’ dispositions toward the Ohio Teacher Evaluation System (OTES). Teacher dispositions are believed to effect performance measures and influence classroom behavior (Rushton, Morgan & Richards, 2007). When teachers perceive evaluation as an accurate reflection of performance and feel valued throughout the process, they are more likely to use evaluation outcomes to improve instructional practice (Tornero & Taut, 2011).

This study was conducted with teachers employed in eight districts located in Miami County, Ohio. Collectively, these districts employed 895 teachers and all were part of the defined study population. The respondents completed a survey constructed by this researcher (Appendix D) through an online survey instrument in the spring of 2015. One hundred and forty-two teachers (15.8%) completed the survey.

Summary of Findings

Demographic profile. Among the 142 teachers who completed the survey; 79% were females. In 2012, 76% of teachers nationally and 75% of teacher in Ohio teachers were females (National Center for Education Statistics, 2012). Thus, respondents in this study were proportionally representative of the gender of national and state populations.
The average (mean) level of respondent teaching experience was 14.3 years. This figure is slightly above the national average, which has been reported to be approximately 13.5 years (NCES, 2008; National Commission on Teaching and America’s Future, 2010). With respect to level of teaching assignment, 57% of the respondents were employed in secondary schools (middle and high schools spanning Grades 6 to 12). In 2011, 51% of teachers nationally were assigned to Grade K-5 (NCES, 2012); therefore, the representation of secondary teachers in this study was slightly higher than the representation of secondary teachers nationally.

Findings pertaining to the research questions. The summary of findings is presented in relation to the four research questions guiding this study.

1. What were the respondents’ dispositions toward specific aspects of OTES?

Respondents stated their levels of agreement with 17 statements pertaining to OTES; they did so by selecting a response from a five-point Likert scale (strongly disagree, disagree, uncertain, agree and strongly agree). Some statements were keyed positively and others negatively in an effort to reduce response bias. A mean response score was then calculated for each statement to determine the respondents’ collective disposition toward each OTES statement. The dispositions were categorized as: positive, moderately positive, neither positive nor negative, moderately negative or negative as described in the two previous chapters.

Overall, the collective dispositions toward the 17 OTES statements were more negative than positive. Specifically, no disposition was categorized as positive; four were categorized as moderately positive; one was categorized as neither positive nor negative; eight were categorized as moderately negative; four were categorized as negative. The
three most positive dispositions pertained to the effectiveness of individualized staff
development over school-wide development, principal compliance with OTES, and the
use of walk-through observations conducted by principals or another evaluator. The
three most negative dispositions pertained to the amount of time required to spend on
OTES using student value-added measures to evaluate teachers, and the use of vendor
assessment scores to evaluate teachers. A full list of means and disposition
categorizations is found in Table 21 in the previous chapter.

2. What were the individual respondents’ overall dispositions toward OTES?

In order to determine a respondent’s personal overall disposition toward OTES a
total score was calculated for his or her aggregate responses to the 17 statements. The
score was then categorized as being positive, moderately positive, moderately negative, or
negative. Respondents’ overall dispositions toward OTES were more negative than
positive. Specifically, 107 of the 142 respondents (75.3%) had a moderately negative
disposition and fifteen (10.6%) had a negative disposition. Only 20 (14.1%) had a
moderately positive disposition and none had a positive disposition.

3. What was the level of association between the criterion variable (teacher
overall personal disposition toward OTES) and each of three predictor
variables (gender, years of teaching experience, and grade level assignment)?

A correlation coefficient was calculated (Pearson product-moment r) to determine
levels of association between the criterion variable and each of three predictor variables.
The coefficients were (a) \( r = -.01 \) for gender, (b) \( r = -.16 \) for years of teaching
experience, and (c) \( r = -.19 \) for grade level teaching assignment. Thus, each of the three
predictor variables had only a small negative association with the criterion variable.
4. *To what extend did the three predictor variables collectively account for variability in the criterion variable?*

A coefficient of determination was calculated to answer the last research question. The multiple, correlation coefficient ($R$) was .256, the value of $R^2$ was .065, and the adjusted $R^2$ was .045. Accordingly, the three demographic variables collectively accounted for only 4.5% of the variability in dispositions.

**Conclusions**

The espoused purpose of evaluation is to improve instruction, which in turn should improve student achievement. Teachers’ dispositions toward OTES were relevant at the time of this study because they almost certainly influenced support for and compliance with OTES. A teacher’s personal overall disposition toward OTES is likely affected by a broader description toward performance evaluation generally and by OTES requirements specifically. In this study, 86% of the respondents were found to have a negative or moderately negative disposition toward OTES. This finding should be considered in light of the fact that teachers generally have not had a positive attitude about the positive effects of performance evaluation. A study conducted by Duffet and associates (2008), for example, reported that only 26% of the teachers they studied had a positive disposition toward the process per se. Because performance evaluation is a high stakes, high-anxiety experience, policymakers and administrators should anticipate teacher skepticism if not covert resistance (Zimmerman & Deckert-Pelton, 2003). Several studies (e.g., Bastarache, 2000; Natriello, 1984; Newton & Braithwaite, 1988) have reported substantial teacher skepticism about the value of their performance evaluation experience.
If a state mandated evaluation program, such as OTES, is not widely-supported by educators it is logical to assume the likelihood of improving instruction is limited. The pervasive negative dispositions toward OTES in this study are largely congruent with findings from previous research studies on teacher opinions about performance evaluation. As examples, teachers in the Chicago Public Schools, with union support, have not adopted changes initiated in their district (Steiner, 2009). Teachers in Minnesota, South Carolina and the Denver Public School system also have resisted changes in their performance evaluation systems even though the stated goals were consistently to improve instruction and provide professional growth opportunities (Steiner, 2009).

Some aspects of OTES have not been normative for Ohio teachers in the past and therefore, this study also examined dispositions toward the model’s specific requirements. The three most negative dispositions related to time requirements and evaluation criteria. With respect to the former, 88.8% of the respondents considered the time they had to devote to the process was excessive. Here again, however, similar negative views about time requirements have been expressed in other studies examining performance evaluation generally or in the context of state mandated systems specifically (e.g., Calabrese, et al., 2004; Kimball, 2003; Steiner, 2009).

At the time of this study, 40 states, including Ohio, were using some form of value-added model to evaluate teachers (Collins & Amrein-Beardsley, 2014). Opposition to using student growth measures, however, was found to be especially negative in this study. OTES requires or permits three such measures: student value-added scores, vendor assessment scores, and SLOs (teacher designed assessments). Here dispositions
toward the first two types of assessments were negative and dispositions toward the third were somewhat negative. A study by Gill et al. (2013) also found that teachers had a more favorable view of SLOs, than of value-added scores even though the reliability of SLOs had not been established. In part, this finding may be explained by the proclivity for educators to be more supportive of requirements that allow them to exercise professional discretion (Kowalski et al., 2008). Nationally, opposition to basing teacher performance on student performance has been substantial among educators (Toch & Rothman, 2008). Thus, findings here regarding value-added and vendor assessments are not unexpected. In fact, several eminent scholars (e.g., Berliner, 2014; Darling-Hammond, 2012) have concluded that basing teacher evaluations on student measures is unreliable, unstable, and biased.

Conversely, the three most positive dispositions toward OTES were related to individualized professional development, principal compliance with OTES, and walkthrough observations. Nearly two-thirds of the respondents, 62%, had positive dispositions toward individualized staff development. Furthermore, 69.7% of the respondents had positive dispositions toward the annual growth or improvement plan requirement. Several studies examining this issue (e.g., Fenwick, 2004; Glickman, et al., 2009) found that teachers had a positive attitude about individualized approaches to professional growth, in part because they viewed the process as motivating, relevant, and self-directed and in part because they had unfavorable views of traditional staff development models.

In this study, 66.25% of respondents believed their principal applied OTES correctly. Although this disposition was somewhat positive, the disposition toward
OTES reducing evaluator bias was somewhat negative. These findings reflect the fact that following policies and guidelines does not de facto reduce subjectivity. The finding on proper implementation is congruent with Zimmerman and Deckert-Pelton (2003) who found that nearly two-thirds of teachers they studied also thought that principals had implemented evaluation policies correctly. The finding on evaluator bias supports earlier research (e.g., Donaldson, 2012) reporting that a majority of teachers believe principals’ evaluative judgments have been biased.

The requirement of walk-through observations also was viewed more positively than most other aspects of OTES. Specifically, 71.2% of the respondents had a favorable disposition toward this requirement. Generally, teacher attitudes about classroom observations have not been positive, largely because of concerns that observers focus on easily measured behaviors rather than evaluating appropriate levels of content, pedagogy, and student learning (e.g., Marshall, 2005; Marzano, 2007; Marzano, 2012; Williams et al., 2002; Youngs, 2013). Perhaps the disposition expressed toward walk-through observations reflects positive opinions toward standardizing observations and toward requiring principals to provide feedback.

All three predictor variables (gender, years of teaching experience, and level of assignment) had a low level of association with the criterion variable (overall personal disposition toward OTES). This outcome suggests that teacher dispositions may be highly associated with more complex variables, such as school culture and workplace socialization. Although personal demographic variables, such as those examined here, are often used to study educator beliefs and behavior, they often have been found to have a low level of association with these criterion variables. For example, Hughes (2006)
found the associations between the three predictor variables examined in this study and two predictor variables (*perceptions of evaluation practices* and *job satisfaction*) not to be statistically significant. Similarly, Schulte and associates (2004) found no statistically significant association between teacher dispositions and predictor variables such as those examined in this study.

**Recommendations**

**Policy.** Logically, a state mandated evaluation policy that is not widely supported by those most directly affected is unlikely to accomplish its intended goals. Teachers may comply with mandates, but mandates often do not alter beliefs and values (Fullan, 2011). The negative dispositions reported in this study, especially in relation to using student data to evaluate teachers, are a matter that should be reconsidered. In fact, following the first year of full implementation (i.e., the period during which this study was conducted), state officials lowered the extent to which student assessments affect teacher evaluations. Furthermore, in September 2015 House Bill 63 eliminated the use of value-added data for the 2014-15 and 2015-16 school years unless a district had a memorandum of understanding to use the data. This provision is due to the transition to a new state testing system. Even so, the legislature did not choose to eliminate value added data nor the substantial weight given to these assessments in subsequent years. As long as student assessments are an integral part of teacher evaluation, determining the validity and reliability of these measures should be a high priority.

**Future research.** The lack of an empirical database on teachers’ dispositions toward OTES was an impetus for conducting this research. The findings here prompt the following recommendations for additional research:
1. The results of the study were impacted by the return rate of the survey, $N=142$ (15.8%). Future research should consider face-to-face administration of the survey, or if using an online survey, extending the duration of the survey’s availability and providing more frequent reminders to increase the return rate. If the return rate were larger, the statistical power would increase.

2. The results of this study could have been impacted by the nature of the study population, which did not include any teachers from large, urban school districts. Therefore, a similar study should be replicated with teachers in large districts.

3. The results of this study could have been impacted by the teachers’ lack of experience with OTES. Since this is the first year of implementation, additional research should be conducted in subsequent years to determine if teachers’ dispositions change with more OTES experience.

4. Qualitative studies should be conducted to gain in-depth and relevant information regarding teacher dispositions toward OTES, especially as it relates to the amount of time teachers invest in the process, the use of student growth measures, and the presence of principal or evaluator bias.

5. This study examined three, predictor variables: gender, years of teaching experience, and grade level assignment. Other possible predictor variables could include teachers’ level of self-efficacy, teachers’ prevailing dispositions toward any type of performance evaluation, teachers’ dispositions toward state mandates in general, and teachers’ belief that performance evaluation will affect student achievement.
6. Given the controversial nature of state mandated evaluation programs, there is a need to determine the extent to which they achieve their two primary objectives; improving instruction and elevating student learning. Empirical data should be the primary determinant regarding the future of OTES.


Chesterfield County Public School System. (2012). *Chesterfield County Public Schools*
Comprehensive Plan. Retrieved from
http://mychesterfieldschools.com/.../ComprehensivePlan%20DesignforExcellence

Retrieved from ProQuest Dissertation and Theses Database, UMI Dissertations
and Publishing. (DP18604)


Correlation analysis for the behavioral sciences, 3rd ed. New York, NY:
Routledge.

the map. Teachers College Record, 116(1). Retrieved from
http://www.tcrecord.org/Content.asp?/ContentID=17291

Council of Chief State School Officer. (2011). Interstate teacher assessment and
support consortium (InTASC) model core teaching standards: A resource


For Educator Compensation Reform. U.S. Department of Education.

Dana, N. & Yendol-Hoppey, D. (2009). The reflective educator’s guide to classroom


classroom leaders improve student achievement. Bloomington, IN: Solution Tree.

http://222.edweek.org/ew/marketplace/.../publications.html


For Supervision and Curriculum Development.


National Council for Accreditation of Teacher Education. (2002). Education Policies and
Procedures. Retrieved from
http://www.ncate.org/LinkClick.aspx?fileticket=fF1tHSizm8%D&tabid=
training workbook. Retrieved from
http://www.nctq.org/docs/OTES_Training_Workbook_Updated.pdf
National Reading Panel. (2000). Teaching children to read: An evidence-based
assessment of the scientific research literature on reading and its implication for
reading instruction. Retrieved from
evaluation processes: A review of six studies using the theory of evaluation and
authority. National Institute of Education: Washington, DC.
New Teacher Center. (2014). Teaching, empowering, leading, and learning Oregon
(TELL OR) survey research brief: How different educators perceive teaching
conditions. Retrieved from http://www.telloregon.org/resources
can be done? Assessment & Evaluation in Higher Education, 33, 301-314.


from
http://www.spokaneschools.org/.../Teacher_Evaluation_Survey_Results_Presentat
ion_PIUS_OPEN_ENDED

Sporte, S., Stevens, W., Healey, K., Jiang, J., & Hart, H. (2013). Teacher evaluation in
practice – the University of Chicago consortium on Chicago School of Research.
Retrieved from http://studentsmatter.org/wp-

Statistic solutions. (n.d.). Retrieved from
http://statisticssolutions.com/point-biserial-correlation

Steiner, L. (2009). Determining processes that build sustainable teacher accountability
https://www.TQSource.org

Taylor, E., & Tyler, J. (2012). Can teacher evaluation improve teaching? Education Next,
12(4). Retrieved from http://educationnext.org/can-teacher-evaluation-improve-
teaching/

Education. Educator Sector Reports. Washington, DC. Retrieved from
http://educationsector.org/sites/default/files/publications/RushTojudgment_ES_Ja
n08.pdf

System: Perceptions and attributions of teachers who actively refuse to

influence on teacher practice and professional growth in four urban high schools.


Alexandria, VA: Association for Supervision and Curriculum Development.


perceptions of the principal’s role in professional evaluation. *NASSP Bulletin, 87* (636), 28-37
APPENDIX A

Panel of Experts

Mrs. Amy Anwanyu, M.S., Principal, Brookville City Schools
Lesley Evans, Ph.D., Teacher in Residence, University of Dayton
Monica L. Richardson, Ph.D., Principal, Northmont City Schools
James Rowley, Ph.D., Professor Emeritus, University of Dayton
APPENDIX B

Communication Sent to District Superintendents to Gain Permission to Contact Survey Participants

Dear Superintendent:

You are receiving this email because you have been identified as a superintendent in a public school district in Miami County. I am seeking permission to contact your certified teaching staff to participate in a research project for a doctoral dissertation at the University of Dayton.

The subject of measuring and improving teacher effectiveness is a critical issue in education. Teacher performance evaluation is a part of the measurement process currently being reviewed as the first year of the Ohio Teacher Evaluation System concludes. One area of interest is teacher dispositions toward performance evaluation. A review of the literature indicates little is known about the dispositions of teachers toward the evaluation process.

In a few days, with your permission, the participants will receive an email asking them to respond to an electronic survey designed to collect information on teacher evaluation practices and attitudes toward the teacher evaluation process in your district.

No personally identifiable information will be collected during the survey. All results will be kept confidential with the data being stored in a password protected site. A summary of the research will be included in the final dissertation project. The results of the survey will be shared with you upon request.

Please reply to this email to grant permission for your teachers to be contacted. If permission is declined, it will not affect your relationship with the University of Dayton or the researcher.

If you have any questions about the survey, please email BWyandt@woh.rr.com.

Thank you for your help in completing this dissertation project.

Sincerely,

Beth A. Wyandt
APPENDIX C

Communication Sent to Survey Participants

Dear Participant:

You are receiving this email because you have been identified as a teacher in a public school district in Miami County. You are being asked to participate in a research project for a doctoral dissertation at the University of Dayton.

The subject of measuring and improving teacher effectiveness is a critical issue in education. Teacher performance evaluation is a part of the measurement process currently being reviewed as the first year of the Ohio Teacher Evaluation System concludes. One area of interest is teacher dispositions toward performance evaluation. A review of the literature indicates little is known about the dispositions of teachers toward the evaluation process.

In a few days, you will receive another email asking you to respond to an electronic survey designed to collect information on teacher evaluation practices and your attitudes toward the teacher evaluation process in your district.

No personally identifiable information will be collected during the survey. All results will be kept confidential with the data being stored in a password protected site. A summary of the research will be included in the final dissertation project. The results of the survey will be available to you upon request. An opportunity will be provided to you during the survey to request this information.

The return of the survey constitutes your informed consent. If you decline to complete the survey, it will not affect your relationship with the University of Dayton or the researcher.

If you have any questions about the survey, please email BWyandt@woh.rr.com.

Thank you for your help in completing this dissertation project.

Sincerely,

Beth A. Wyandt
APPENDIX D

Teachers’ Dispositions toward OTES Survey

Welcome!

Thank you for agreeing to complete this survey regarding the Ohio Teacher Evaluation System (OTES). The purpose of this survey is to gather teacher dispositions toward OTES to add to the knowledge base on teacher evaluations and to assist policymakers and educational leaders as they make decisions regarding the OTES model. As legally mandated, the law was implemented during the current (2014-15) school term.

Your participation in the study is voluntary. You will not be compensated for your participation. If you choose to participate in this study, you will be asked to complete an online survey consisting of 17 statements with Likert-type response categories. Your time commitment to complete the survey will be approximately 5 minutes. There are no known risks from participating in this research.

If you decide to participate in the study, you will receive a copy of the results from the research for your reference upon request. If you decide not to participate after starting the online survey, you can withdraw from the study by closing your internet browser window. However, once you have completed and submitted your survey data, you cannot withdraw from the study due to the anonymous format of the study. If you decline to complete the survey, it will not affect your relationship with the University of Dayton or the researcher.

No personally identifiable information will be collected during the survey. Any individual data obtained in relation to this research study will remain unidentifiable due to the anonymous nature of the survey. Overall information and results collected because of your participation in the study will be used for the completion of a doctoral dissertation.

Selecting the “next” button at the bottom of the page gives your consent to participate and begin the survey.

Section I. Elements of OTES (This section includes statements pertaining to specific aspects of OTES).

Please indicate your level of agreement with each statement by selecting one of the five response options: Strongly Disagree, Disagree, Uncertain, Agree, and Strongly Agree.
1. Requiring teachers to develop annual growth plans is unreasonable.
   
   Strongly Disagree
   Disagree
   Uncertain
   Agree
   Strongly Agree

2. Individualized staff development is more effective than school-wide development.
   
   Strongly Disagree
   Disagree
   Uncertain
   Agree
   Strongly Agree

3. Student value-added measures should be used to evaluate a teacher’s performance.
   
   Strongly Disagree
   Disagree
   Uncertain
   Agree
   Strongly Agree

4. Student learning objectives (SLOs) should not be used to evaluate a teacher’s performance.
   
   Strongly Disagree
   Disagree
   Uncertain
   Agree
Strongly Agree

5. Vendor assessment scores should be used to evaluate a teacher’s performance.
   Strongly Disagree
   Disagree
   Uncertain
   Agree
   Strongly Agree

6. The approved alternate framework (42.5% teacher performance; 42.5% student growth; and 15% additional measures) should not be used to evaluate a teacher’s performance.
   Strongly Disagree
   Disagree
   Uncertain
   Agree
   Strongly Agree

7. The required level of conferencing between a teacher and the principal (or other evaluator) is excessive.
   Strongly Disagree
   Disagree
   Uncertain
   Agree
   Strongly Agree

8. Walkthroughs by the principal (or other evaluator) should be used to evaluate teacher performance.
   Strongly Disagree
   Disagree
   Uncertain
Section II. Experiences with OTES (This section includes statements pertaining to your experience with OTES this year).

9. Instructions for gathering additional teacher performance evidence were clear.
   - Strongly Disagree
   - Disagree
   - Uncertain
   - Agree
   - Strongly Agree

10. Compared to previous evaluation methods, evaluator bias was less common because of OTES.
    - Strongly Disagree
    - Disagree
    - Uncertain
    - Agree
    - Strongly Agree

11. The principal (or other evaluator) followed the OTES instructions correctly.
    - Strongly Disagree
    - Disagree
    - Uncertain
    - Agree
    - Strongly Agree

12. OTES increased the accuracy of evaluations.
    - Strongly Disagree
    - Disagree
Uncertain
Agree
Strongly Agree

13. OTES had a negative effect on principal (or other evaluator)-teacher relationships.
   Strongly Disagree
   Disagree
   Uncertain
   Agree
   Strongly Agree

14. Staff development was more individualized because of OTES.
   Strongly Disagree
   Disagree
   Uncertain
   Agree
   Strongly Agree

15. The principal (or other evaluator) assisted teachers to develop individual growth plans.
   Strongly Disagree
   Disagree
   Uncertain
   Agree
   Strongly Agree

16. The level of evaluator subjectivity under OTES was low.
   Strongly Disagree
   Disagree
119

Uncertain
Agree
Strongly Agree

17. The amount of time teachers had to devote to OTES was excessive.

Strongly Disagree
Disagree
Uncertain
Agree
Strongly Agree

Section III: Demographic Information

18. Gender
   a. Female
   b. Male

19. Years of teaching experience (including current year)
    [Pull down screen with options of 1 to 50]
   a. 1

20. Current level of assignment
    a. Grades K-2
    b. Grades 3-5
    c. Grades 6-8
    d. Grade 9-12