Effects of Teacher Evaluation on Teacher Job Satisfaction in Ohio

Pamela R. Downing
The University of Findlay

Abstract

The purpose of this quantitative study was to explore whether or not increased accountability measures found in the Ohio Teacher Evaluation System (OTES) impacted teacher job satisfaction. Student growth measures required by the OTES increased teacher accountability. Today, teachers are largely evaluated based on the results of what they do in the classroom as measured by student performance, rather than what they do to accomplish the task of teaching. Understanding the impact of the OTES on teacher job satisfaction is important as a connection between teacher job satisfaction and quality education has been identified.

Keywords: accountability, job satisfaction, teacher evaluation, student growth measure, high-stakes testing

Introduction

In an effort to improve educational outcomes for Ohio's K-12 students, Ohio's legislators have mandated through the legislative process, rapid change and increased accountability measures causing Ohio's educators to experience change within the profession at an unprecedented rate (Achieve, Inc., 2007). Currently, the state of Ohio's legislators have written and continue to write and vote for a wide variety of educational policy changes, especially policies increasing accountability measures for educators in K-12 educational settings. These policy changes and increased accountability measures have a myriad of compelling ramifications for the state's teachers. The impact of increased accountability measures imposed upon Ohio educators provides a developing opportunity to understand if and how increased accountability affects teacher job satisfaction.

As an illustration of increased accountability and policy changes, several pieces of legislation stand out. One of the legislative pieces was House Bill 555. Governor Kasich signed the bill into law in December 2012, and it became effective in March 2013. This legislation removed previous designations for Ohio K-12 schools such as "continuous improvement" and "academic watch" and replaced it with an A-F grading system. This new grading system for Ohio schools was based on student academic progress and teacher performance, determined through a new and rigorous evaluation model known as the Ohio Teacher Evaluation System (OTES). Another legislative piece, House Bill 153, signed into law in 2011, spelled out the new teacher evaluation system, which combined student growth measures with observable performance ratings of the teacher to provide a holistic determination of a teacher's designation (Ohio Department of Education, 2011). A teacher may be designated as Accomplished, Skilled, Developing, or Ineffective (Ohio Department of Education, 2015b).

Additionally, the Ohio General Assembly passed Senate Bill 316 into law in June 2012. This legislation included the new third grade reading guarantee. At present, Ohio schools are graded on the number of students reading at grade level. While all students in K-3 receive interventions if they are not found to be reading at grade level, only third grade students face retention based on the Ohio Achievement Assessment for Reading. A student must receive a score of 400 or higher to be considered to be reading at grade level. This policy puts third grade teachers in the spotlight, potentially adding stress and increased pressure to perform. Not only will poor student performance on the Ohio Achievement Assessment for Reading negatively impact a teacher's evaluation, there are financial implications for a school district if a student is not reading at grade level and is retained. For example, a series of interventions must be put into place for that student. These interventions may include increased time devoted to reading instruction for that student and intervention services from outside providers (OCTELA, 2012).

More notable legislative changes, found in Senate Bill 165 were college and career readiness standards that change how Ohio's students are assessed prior to graduation. Because schools will be judged on their preparation of high school students to be college and career ready, this policy change impacts how Ohio students will be tested. For more than a decade, students have been required to take the Ohio Graduation Test (OGT). Beginning in 2015, Ohio students began taking a series of PARCC (Partnership for Assessment of College and Career) Assessments and "end-of-course" exams provided by AIR (American Institute for Research) Assessments. However, legislation was passed with the 2015-2016 budget by the General Assembly and signed by Governor Kasich ordering the Ohio Department of Education to discontinue the use of the PARCC Assessments. Instead, the AIR tests would be used for all tested subjects (Ohio Department of Education, 2016). An overlap of these assessments was anticipated causing concern among educators, parents, and students that important instructional time would be sacrificed in order to administer these assessments (Guilfoyle, 2006). These aforementioned examples are merely a few of the legislated changes facing Ohio educators. Teachers, both experienced and inexperienced, have complained of being overwhelmed by the pace of the policy changes and by the uncertainty of how they will be evaluated under increased accountability measures (Franco, Zigler, & Lindsey, 2013). Teachers lacked faith in the validity of student growth measures as a component of their evaluation. The reason for this lack of faith is the data reflect on the previous year's student growth rather than the current year for which teachers are receiving the evaluation rating.

Connected to teacher evaluation is job satisfaction. Job satisfaction, in general, may be negatively impacted by factors such as low pay, dwindling resources affecting employee performance, media disparagement, frequent changes in educational policy, and circumscribed accountability measures lacking teacher input (Scott, Stone, & Dinham, 2001; Van den Berg, 2002). Zembylas and Papanastasiou (2005) found that educator job satisfaction is related to teacher empowerment. In fact, they claimed a less than satisfying evaluation system was found to diminish teacher empowerment and therefore job satisfaction. Job satisfaction was associated with teacher motivation, and these occupational attitudes have been linked to quality education in schools (Evans, 2000).

Another crucial issue for educational leaders to consider is that in order to provide effective leadership and supervision to K-12 teachers, it is imperative for administrators to understand how frequent and fast-paced policy changes and increased accountability measures affect teacher job satisfaction. According to Linda Evans (2000), co-director of the University of Warwick's Teacher Development Research and Dissemination Unit, job satisfaction, motivation, and morale are all work-related attitudes. Morale is different from job satisfaction in that morale is future-oriented and anticipatory, while job satisfaction is present-oriented or a response to a situation (Evans, 2000). Evans (2000) defines motivation as "a condition, or the creation of a condition, that encompasses all of those factors that determine the degree of inclination towards engagement in an activity" (p. 179). This study focused on the work attitude, job satisfaction as defined by Fuming and Jiliang (2007): "the degree of satisfaction a worker evinces for the work in which he or she is engaged" (p. 87).

Another important concept to be defined for this study is accountability. Wood and Winston (2005) said "Accountability refers to employees' beliefs about the degree to which they will be required to justify their actions at work to one or more individuals who hold reward or punishment power" (p. 85). In the case of Ohio teachers, accountability will be measured by student growth (e.g., "value-added"), adequate yearly progress (AYP), and student achievement in the form of the Ohio Graduation Test, which was slated for elimination. New AIR assessments or end-of-course exams were being developed and approved by the Ohio Department of Education and were expected to provide a measure of academic achievement. Earlier research suggested increased accountability may stimulate both positive and negative results connected with job satisfaction (Hochwarter, Ferris, Gavin, Perrewe, Hall, & Frink, 2007). This implication is important because a strong correlation between teacher job satisfaction and quality of education has been identified (Persevica, 2011).

Accountability has become a buzzword according to Lerner and Tetlock (1999). Accountability continued to be defined as the "implicit or explicit expectation that one may be called on to justify one's beliefs, feelings, and actions to others" (Tetlock, 1992, p. 332). The term accountability created stress due to the inference that accountability implied that people who did not justify actions satisfactorily would endure adverse consequences (Stenning, 1995). The term accountability used in education-oriented discussions connoted an ethical responsibility of the school or teacher for effective education (Levit, 1972).

This call for accountability, a nationwide movement that was technocratic and efficiency-oriented (Levit, 1972), has grown in momentum and increased in volume. States across the nation, including Ohio, have developed nearly identical programs to ensure educational accountability, with nearly all of the programs utilizing large-scale assessment results (Popham, 2000). The use of large-scale assessments was no minor detail in the world of educators, because the movement of employing accountability systems based on student testing "had the potential to become one of the major reform efforts in American education in this century, perhaps equal in impact to such movements as the development of the comprehensive high school or the racial integration of public education" (Ramirez, 1999, p. 205). While many reforms exist in Ohio to address the call for increased accountability, the current study focuses on one aspect of increased

accountability reform, that being the introduction of the Ohio Teacher Evaluation System.

The OTES was developed to reform the state's educational system and advance student growth and achievement as called for in No Child Left Behind and Race to the Top, both federal attempts to improve the nation's competitive standing in the world economic arena. The OTES was designed by Ohio teachers, administrators, and college and university faculty along with various educational associations under the guidance of the Ohio Educator Standards Board (Ohio Department of Education, 2007). These Ohio educational professionals worked collaboratively with national experts on teacher evaluation (Ohio Department of Education, 2013). The OTES was designed to more closely align teacher evaluation with the new Ohio Standards for Educators. The OTES became effective for the 2013-2014 school year. Ohio Department of Education materials further claimed the OTES was dedicated to teacher growth and student achievement (Ohio Department of Education, 2015b). It sought to strengthen and re-vamp teacher evaluation. Previously, teacher evaluations were often superficial, offering very little valuable feedback to teachers. According to the Ohio Department of Education, these efforts to advance teacher evaluation were for the purpose of seeking improvement in student educational growth. Under the OTES, 50% of a teacher's evaluation was based on student growth measures (e.g., "value-added"). The other 50% was based on a series of formal and informal observations conducted by administrators. These two components were the primary source of legislative debate in Senate Bill 229 (Harris, 2015).

While OTES is relatively new, its future remains unclear. Senate Bill 229 made its way to the House of Representatives where it became House Bill 362. The Ohio General Assembly passed it on June 3, 2014. The highly contested changes reduced the frequency of evaluations for skilled or accomplished teachers and allowed districts to choose between the original evaluation system structure and an alternative structure. The alternative framework for the OTES included a 42.5%, 42.5%, and 15% division of percentages of category weights (Ohio Department of Education, 2014). The alternative framework was once again amended in Ohio House Bill 64 for the 2015-2016 school year. Changes included making the teacher performance rating worth 50% of the evaluation and student growth worth 42.5%. The alternative component chosen by the district was given a value of 15% (Ohio Department of Education, 2015a). Due to the short interval of time since OTES was first implemented, its effect on student growth and achievement remains unclear.

Rationale and Significance of the Study

Shifting the teacher evaluation process in Ohio from looking at what teachers do in the classroom to what students learn was a major alteration in the teacher evaluation paradigm. This deviation from traditional teacher evaluations to evaluations that include student test scores as a measure of teacher effectiveness creates consequences in the teaching profession. Using student test scores to determine teacher effectiveness as part of the teacher evaluation escalates the accountability element. Because the OTES was relatively new, little to no research exists, creating the opportunity to conduct an investigation. In light of the fact that legislators continue to create and tweak educational

policies regarding teacher evaluation, this study seemed significant and may provide important information for consideration in policy development at the state level. Moreover, educational administrators may find the results of the current study helpful as they employ the state-level teacher evaluation policies in their local districts.

Purpose of Study

Because student test scores became part of the calculation in teacher evaluation, teachers may have perceived the OTES as holding them more accountable for what happened in the classroom. Accountability may impact teacher job satisfaction. As mentioned earlier, positive correlations between teacher job satisfaction and the quality of education in the classroom have been identified (Persevica, 2011). Persevica (2011) concludes that teacher job satisfaction is a fundamental element of quality of education. This exploration seeks to gain insight into the relationship between accountability and teacher job satisfaction, specifically in Ohio K-12 public education. Therefore, the purpose of this study is to determine the impact of increased accountability conveyed through the Ohio Teacher Evaluation System on teacher job satisfaction.

Research Questions

In order to ascertain the outcomes related to teacher job satisfaction created by the various elements of the OTES and the perceived increased accountability, the researcher sought to determine answers to the following questions:

1. Is OTES associated with an impact on teacher job satisfaction?
    a. If so, is the relationship positive or negative?
2. Which components of OTES, if any, are most associated with teacher job satisfaction?

Theoretical Framework

The driving question pertaining to the relationship between accountability and teacher job satisfaction lead to the study hypothesis: increased accountability, perceived in the various elements of the OTES, has a positive rather than a negative impact on teacher job satisfaction. This hypothesis was based on Maslow's Hierarchy of Needs, specifically focusing on the top portions of the Maslow pyramid: esteem needs and self-actualization needs. The esteem needs were based on respect of others and respect by others, as well as self-esteem and achievement (Maslow, 1943). Through increased accountability, the successful teacher would potentially feel greater respect and heightened accomplishment. Maslow described self-actualization needs as incorporating morality and creativity among other elements (Maslow, 1943). Successful teachers would likely feel a moral obligation to help students achieve in their classrooms. Therefore, in an effort to reach all students and meet the specific learning needs of each child, teachers were apt to express creativity in their instructional design.

Delimitations

The study was designed to investigate the impact of increased accountability, via components of the OTES, on teacher job satisfaction. Because the researcher was investigating components of the OTES, and not evaluation systems in general, the current study was focused on K-12 public education teachers in the state of Ohio. Participation in the study was for teachers who have been evaluated through the OTES. Parochial and charter school instructors were not included in the study as they were not required to adhere to the structure of the OTES.

Limitations

Identified limitations for this study include: First, the study was based on a convenience sample rather than a random sample. Because teacher e-mail addresses were unattainable, the survey link was sent to superintendents and principals to forward to their instructional staff. The majority of districts receiving the survey link were in northwestern Ohio, where the primary investigator has the most professional contacts. Second, the researcher has no way to know how many teachers received the survey link to determine a response rate. Some superintendents or principals may have failed to see the e-mail with the survey link or may have determined they did not want their teachers to participate. Conversely, some teachers may have forwarded the survey link to peers in other districts. Third, section two of the survey incorporated questions from the P. E. Lester Job Satisfaction Questionnaire (TJSQ). The researcher inadvertently left off a question from the supervision section of questions. While the question was not critical to the current research, the researcher carefully analyzed the statistical results for all supervision questions. The TJSQ item omitted from the survey was "My immediate supervisor treats everyone equitably" (Lester, 1982, p. 13). The fourth and final limitation affected reliability. The survey included only one question each pertaining to student growth, pre-conference and post-conference, to link to job satisfaction. Therefore, reliability of the survey was decreased.

Researcher Bias

With the introduction of the new Ohio Teacher Evaluation System (OTES) teachers were faced with a presumably more rigorous evaluation system that included two components. One component was based on teacher performance determined through both formal and informal observations conducted by a supervisor. The other component as previously mentioned, was a student growth measure. Teachers were now to be evaluated based on student performance, in terms of measuring student growth, in addition to their own performance in the classroom (Ohio Department of Education, 2013). The researcher chose this topic out of curiosity regarding whether teachers would experience an increase or decrease in job satisfaction due to the increased accountability elements of the OTES.

The researcher entered the study with bias based on the assumption that the OTES increased accountability and was further biased in the expectation to learn that most teachers had improved job satisfaction levels due to increased communication with their

evaluator and improved methods for tracking student growth. These two elements were byproducts of the increased accountability measures found in the OTES. Researcher bias favored teachers' desires to be effective and against critics who claimed educators were not concerned with having a positive effect on student learning. The researcher anticipated the study would indicate that increased accountability improves teacher job satisfaction with the expectation this exploration would provide some insight including evidence of the contrary and did not feel compelled to deny any evidence discounting her opinion. The researcher was careful to prevent bias from impacting the study's results. Preventative measures were taken, such as asking proofreaders to look for instances of potential bias.

Review of the Literature

The state of Ohio's General Assembly imposed a stream of increased accountability measures for educators in K-12 educational settings. These accountability measures directly impacted teachers throughout the state. This study examines the impact of increased accountability measures on teacher job satisfaction. Research of the literature conducted in preparation for the current study implies that increased accountability may stimulate both positive and negative results connected with job satisfaction and correlated teacher job satisfaction and quality of education (Hall, Zinko, Perryman, & Ferris, 2009). Understanding the impact of increased accountability on teacher job satisfaction will be beneficial to policy makers and educational leaders as they make decisions regarding teacher accountability measures in the future.

Educators and policy makers alike acknowledged what astute parents figured out a long time ago: the competency of the individual teacher counts (Danielson, 2001). The accountability movement, which began with the 1958 National Defense Education Act, led to the evolution of the standards-based reform movement and culminated with The Ohio Teacher Evaluation System (OTES) as a response to the accountability movement in the state of Ohio. Political aspects of the accountability movement and how legislators benefited from the perpetuation of the belief that schools are failing are pertinent to the study of accountability. Additionally, the pros and cons of high-stakes assessment, value added measures, and the validity of the data are important elements for a holistic study of teacher accountability, evaluation, and job satisfaction. Issues of high-stakes testing and how such assessments positively and negatively impact education are key components of the research. The progression of teacher evaluation will be traced, as many of the transformative elements were found in the OTES.

Undeniably, teacher evaluation has been the primary tool for increasing the accountability of job performance for teachers. Procedures for the evaluation of teachers are typically spelled out in contracts as determined by the school district and the local bargaining agent. A number of states have legislatively mandated teacher evaluation (Stodolsky, 1984). Ohio was one of those states. External stakeholders were bolstered by teacher evaluations that reflected the success of the school (Peterson, 2004). However, it was not until the turn of the 21st century that the focus of teacher supervision turned to the evaluation process (Marzano, Frontier & Livingston, 2011).

Initially, teacher evaluation was most commonly made up of an anecdotal report, summarized judgment, numerical rating, or checklist completed by the school principal

after visiting the classroom one or two times during the school year (Boyd, 1989; Loup, Garland, Ellett, & Rugutt, 1996; Stodolsky, 1984). Teachers were evaluated on appearance and personality, in addition to rapport with students, preparation for teaching, content knowledge, classroom management, and professional contributions (Stodolsky, 1984). Teacher evaluation typically lacked common ideals and antecedents regarding what constituted good teaching, not to mention providing insufficient feedback for teachers (Danielson & McGreal, 2000). In developing a framework for teacher evaluation, Danielson and McGreal (2000) suggest teacher evaluation be based on a set of teaching standards. They further recommend the focus of teacher evaluation be formative in nature.

According to Stodolsky (1984), the format of direct observation for teacher evaluation was limited in the information it could produce. Moreover, the process of teacher evaluation had an issue with low validity (Medley & Coker, 1987). By the late 2000's the practice of teacher evaluation came under scrutiny (Marzano, et al., 2011). Indeed, Peterson (2000) maintains that evaluations are not useful in terms of improving instruction. As mentioned earlier, in 1987 the NBPTS created standards for teachers which included the following components: identified and defined elements of good teaching, a rubric outlining levels of performance, more frequent observations and collection of artifacts for a more holistic evaluation, and training for the evaluator (Danielson & McGreal, 2000).

One response to the scrutiny of teacher evaluation practices was the Measures of Effective Teaching Project (MET Project). The MET Project focused its mission and efforts on determining and disclosing techniques for measuring effective teaching. The MET Project, funded by the Bill and Melinda Gates Foundation, involved in excess of 3,000 teachers whose participation was voluntary. The MET Project aimed to provide tools for teachers to be successful at improving student achievement. The primary goal of the MET Project was to determine how evaluation could be used to develop outstanding teaching. The research conducted was based on research showing that "a teacher's contribution matters more than anything else within the school" (Cantrell & Kane, 2013, p. 1).

Ultimately, the research of the MET Project found various elements made up effective teaching and, therefore, must be evaluated using a variety of measures. Furthermore, researchers established that evaluations must be both valid and reliable in order to be worthwhile (Coker, Medley, & Soar, 1980; Medley & Coker, 1987; Kane & Staiger, 2012; Scriven, 1981; Stodolsky; 1984). The MET Project defines "valid" as teaching measures proven to lead to student learning and defines "reliable" as reflective of a typical performance, without the influence of the observer or the particular group of students (Kane & Staiger, 2012).

In response to the call for accountability, researchers began to consider various data sources for the purpose of teacher evaluation. As intended, elements added to the traditional teacher evaluation brought increasingly more accountability to teachers and their supervisors. For example, Value-Added Measures (VAMs) became a component used for teacher evaluations. VAMs were used to ascertain the impact a teacher had on student growth. Typically, this was determined through a statistical analysis of effectiveness based on standardized test scores. VAMs evolved due to a growing interest

in measuring teacher effectiveness and data-based decision-making (Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012).

According to Darling-Hammond, et al. (2012), issues with VAMs as appropriate measurement of teacher effectiveness were discovered. For instance, VAMs were unpredictable. Furthermore, the value-added score of a teacher could change based on the students assigned to their classroom, as there are many elements that contribute to or hinder student growth (Darling-Hammond, et al., 2012; Everson, Feinauer, & Sudweeks, 2013). Value-Added Measures were unable to discern those elements from teacher effect (Darling-Hammond, et al., 2012). Per Everson, et al. (2013), VAMs should not be used singularly for teacher evaluation because measurement problems exist that need to be solved.

Despite arguments about using test scores to make comparisons, state lawmakers were incorporating student growth measures into teacher evaluation systems. Veritably, "it is genuinely difficult to find a large-scale educational assessment that isn't playing some sort of role in a local or regional accountability drama" (Popham, 2000, p. 283). As a result of including student growth measures into teacher evaluation, the focus of teacher evaluation changed from one of inputs to one of outputs (Levit, 1972). The focus of evaluation used to be to examine what teachers do and the tasks they perform. These are inputs. What a student knew and could do with that knowledge were defined as outputs. The new focus on student achievement focuses on outputs (Kellaghan, Stufflebeam, Pearlman, & Tannenbaum, 2003). Student growth measures are calculated in terms of Adequate Yearly Progress (AYP) and also VAMs.

Value-Added Modeling is used to show the effects of the school and teacher-on-student achievement or growth. Such information was useful in reflecting the importance of the teacher in the outcomes of student learning. Value-Added Modeling was an important component of the high-stakes test and accountability movement (McCaffrey, Lockwood, Koretz & Hamilton, 2003). Adequate Yearly Progress was a measurement of teacher and school contributions to a student's learning. It has been an essential tool for holding teachers and schools accountable (Kupermintz, 2003). However, Kellaghan, et al. (2003) argue that test misalignment (when tests do not align with what teachers are asked to teach) would provide unfair results. Moreover, standardized tests typically cover basic recall of information (Kellaghan, et al., 2003) rather than higher-level thinking such as evaluation or analysis. This could penalize teachers who focus on critical thinking, a 21st century skill. Adding to the controversy are teachers' perspectives of standardized achievement tests. According to Urdan and Paris (1994), teachers, by and large, have not respected the validity of the tests. They have had no say in what tests were given or in how the results were used. The results are not useful in determining how to help their students (Urdan & Paris, 1994). Because of this, Urdan and Paris (1994) point out that teachers may have employed methods that undermined students' test score validity.

Half of each teacher's evaluation would have been derived from the degree to which his or her students learned during the school year. Student growth measures were a mechanism for ascertaining the degree of academic gains students made (Ohio Department of Education, 2013). This was done by calculating student growth between two points in time, also referred to as the interval of instruction. Because there was no common assessment shared by all teachers, the student growth element was challenging.

Three methods for measuring student growth were determined. The first method was Value-Added. Value-Added Measures were discussed earlier in the literature review. The OTES required teachers to use Value-Added data, if it existed, for their students. A second method, if legally acceptable, is the district's local student growth measures. If Value-Added data were unavailable, districts were instructed to use assessments referred to as Approved Vendor Assessments. Such assessments were offered by national testing vendors if they were on the approved list for the state of Ohio. The third method for determining student growth was referred to as Locally-Determined Measures. These measures were to be used in areas such as art or music, where Value-Added data and Approved Vendor Assessments were not available (Ohio Department of Education, 2015b).

In situations where the third method was necessary, the districts were charged with creating the opportunities to measure student progress. Districts were able to produce measures through a locally authorized procedure. Locally Determined Measures included Student Learning Objectives (SLOs) or Shared Attribution. An SLO was one way to establish a teacher's influence on student learning. Student Learning Objectives were intended to cover a long-term interval of instruction and included a target for academic growth for each student. As previously stated, shared attribution was another Locally Determined Measure. Shared attribution was a growth measure that was attributed to or shared by a group. Shared attribution was helpful for supporting collaboration for meeting school goals (Ohio Department of Education, 2015b).

Without a doubt, excellence in instruction goes much further than test scores, such as provoking a love for learning, developing critical thinking skills, and encouraging creative thinking (Everson, et al., 2013). A drawback of VAMs was that they were comparative, especially when they were used to make employment decisions regarding retention and promotion. This practice has pitted teachers against each other instead of encouraging an environment of cooperation and development of professional learning communities (Everson, et al., 2013).

Using student achievement data for the evaluation of teachers was a problematic undertaking (Stronge & Tucker, 2000). Effective teachers possessed a variety of strengths which they brought to the classroom. What makes one teacher effective is potentially different from what makes another teacher effective. Because teachers are individuals, their evaluations ought to have been tailored or differentiated (Peterson, 2004). Additionally, the practice of instruction was both methodical and impromptu (Stodolsky, 1984). Weiss and Weiss (1998) recommend that "teaching needs to be understood dynamically in its multiple contexts, and performance data needs to be gathered from diverse sources" (p. 4). Such factors are important to consider during the evaluation process for an improved understanding of results.

Teachers were both apprehensive and dubious of the evaluation process (Peterson, 2000). The OTES was a new and unknown entity and, according to Peterson (2004), "Teachers [would] not support systems with inadequate procedures and components" (p. 63). Teacher evaluation systems were relevant when they focused on aspects of teaching that were seen as valuable to both the evaluator and the teacher (Iwanicki, 2001). In other words, evaluation systems must make sense to the practitioner (Peterson, 2004). Teachers were concerned about factors related to the nature of teaching itself (Danielson, 2001). Evaluation has been viewed as an activity in which the teacher participates and

also one that encourages reflection on the part of the teacher. Such components made the process more meaningful to teachers (Weiss & Weiss, 1998). Unfortunately, teacher evaluation often encouraged teachers to follow procedure rather than actually advancing teacher performance (Johnson, 1990). Placing an emphasis on procedure was an ineffective means to achieve the desired goal of academia, which is student learning. Lynn (2013) states that teacher views should be taken into account when creating an evaluation system. According to Lynn (2013), "Teachers viewed evaluations as a tool for improvement, while school reform advocates and some parents viewed evaluations as a way to dismiss teachers who were not performing well enough" (p. 208). While the accountability movement called for appraisals based on standards and student growth, the goal of evaluation should have been to develop systems to increase productivity of the school, not systems to fire people (Iwanicki, 2001). Inarguably, when a teacher is consistently ineffective, they should receive more comprehensive evaluation with the possibility of termination (Iwanicki, 2001).

Standards-based accountability has been one of the most important accomplishments of the reform movement, as widely accepted standards for the teaching profession (clearly defining what a teacher should know and be able to do) were endorsed (Kellaghan, et al., 2003). Evaluating teachers based on the standards was the logical next step toward accountability. The purpose of standards-based teacher evaluation systems is to provide standards and rubrics in order to determine the effectiveness of instructional choices and to provide accountability (Borman & Kimball, 2005). In fact, the focal point of legislative policy-making at both the state and federal levels has been the standards movement (Seashore Louis, Febey, & Schroeder, 2005).

Many elements of the OTES are directly connected with both accountability and teacher job satisfaction. While a direct link between teacher job satisfaction and student achievement is a subject of disagreement between researchers, its overall importance in the realm of educating students is conclusive. When the research for this study began, the state of Ohio was in the process of fully implementing the OTES. A look at OTES components and their impact on teacher job satisfaction was needed. Ohio's teachers face multiple challenges and stressors in their work. How did teacher job satisfaction fare in the aftermath of the OTES?

Methodology

As stated previously, the OTES is fairly new for Ohio's educators. Its impact on education for Ohio's youth has not yet been studied. However, a critical connection between teacher evaluation and teacher job satisfaction has been made, as has the relationship between teacher job satisfaction and quality of education. Therefore, the researcher determined an investigation into the topic was necessary.

This study used primary data collected through a quantitative, non-experimental research design. Non-experimental research was identified as appropriate for the study as the researcher sought to understand the dependence of variables through correlations. The researcher collected primary data using an electronic survey that included three sections. The three sections included demographic, job satisfaction, and OTES-specific questions. Subjects of the research were Ohio K-12 teachers who were evaluated through the OTES format. A digital survey was chosen for conducting research as it required

minimal effort with the potential for expedient results. The online survey was sent to superintendents and principals, selected by the researcher, who forwarded the survey to teachers. The online survey was open for data collection from May 20, 2015 until August 31, 2015. Mid-May was the target for sending the survey as teachers received their written evaluations in May. This timing would allow respondents to complete the survey while the evaluation experience was fresh in their minds. Data were analyzed to determine if increased accountability elements of the OTES positively or negatively impacted teacher job satisfaction.

Reliability for survey items regarding teacher job satisfaction was previously established by P. E. Lester, the creator of the TJSQ. Internal consistency of the TJSQ was established through calculation of an Alpha coefficient. "The total scale for the sample (N=526) was .93" (Lester, 1982, p. 2). The P. E. Lester TJSQ was chosen for use in the current study as it was designed to be used with teachers and educational research with language specific to the field. Additionally, the TJSQ already had established rates of reliability and validity. When Lester (1982) tested for reliability, she did so for the total and for each of nine factors including Supervision, Colleagues, Working Conditions, Pay, Responsibility, Advancement, Security, and Recognition. Lester's work allowed the researcher to break up the questionnaire by subscales or factors, using only four of the factors in the current study, and still retain the established reliability. The four factors from the Lester instrument used in the current study included Factor 1: Supervision; Factor 5: Responsibility; Factor 8: Security; and Factor 9: Recognition (Lester, 1982).

Lester established content validity by having the job satisfaction questionnaire examined by experts in the field. Moreover, the instrument was examined for the plan and procedures used to construct the instrument. The instrument was also examined for how instructions were written, how the items on the instrument were ordered, and which items were chosen to be included in the questionnaire. A modified Q sort was used to achieve content validation. Any item with less than 80% agreement was either rewritten or rejected by Lester. In order for Lester to generate an amalgamation of 120 items, each potential item was analyzed for its length, its clarity and repetitiveness, and particularity to the field of education (Lester, 1982).

Validity for the OTES Impressions section of the survey was established through use of expert analysis of the items. Each item in the third section of the survey, referred to as the OTES Impressions section, was deemed necessary and important to elicit and establish a thorough snapshot of participants' perceptions of the OTES. All items were piloted with 27 teachers for clarity, and the entire survey was piloted for the length of time a potential participant might expect to spend to complete the survey. Seventeen teachers out of 27 submitted completed surveys. The median time spent taking the survey was 6.5 minutes. Ambiguous items were refined for precision.

Results

The study's target population comprised licensed K-12 teachers in public education in Ohio who had been evaluated under the Ohio Teacher Evaluation System. All participants received the survey with an invitation to participate, which included a statement of implied consent, in addition to contact information for the researcher and the survey. The researcher received 321 survey responses. Thirty-one of the surveys were

incomplete; some respondents chose to skip questions within the survey but otherwise submitted a completed survey. The researcher, therefore, received 290 fully completed survey responses, which far exceeded the anticipated 100 responses. This was a 90% completion rate. For those surveys that were incomplete, the respondents most frequently stopped answering questions at the end of the demographic response section. Despite the fact that the survey took less than ten minutes on average, the appearance of length seems to have been a limitation.

Quantitative data were analyzed using Pearson product-moment correlations to determine if relationships between the OTES, in addition to various components of the OTES, and teacher job satisfaction held any significance. In order to answer the research question, "Is the OTES associated with an impact on teacher job satisfaction?" the researcher used OTES Impression and OTES Performance to delineate the OTES. A numerical value was designated to each of the OTES performance ratings, which allowed the researcher to create a new variable labeled OTES Performance. OTES Impression data was derived from section 3 of the survey. Section 3 of the survey was written in a Likert-scale style with a numerical value appointed to each response. Values were averaged to provide an OTES Impression variable.

Once the researcher had the OTES Impression and OTES Performance variables, a job satisfaction score was sought. Job satisfaction data were derived from section 2 of the survey, which included questions from the P. E. Lester's Teacher Job Satisfaction Questionnaire (1982). A numerical index was calculated for the purpose of demonstrating the relationship between the variables. A correlation matrix was then created using Pearson product-moment correlation coefficients. In response to research question one, "Is the OTES associated with an impact on teacher job satisfaction?" the researcher identified a weak-to-no relationship between OTES performance and overall teacher job satisfaction. In considering the relationship between OTES performance and the sub scores of teacher job satisfaction, the researcher found a weak-to-no relationship between Sub-score 1, Supervision, and teacher job satisfaction. The relationship between OTES Performance and Responsibility (Sub-score 2), was also a weak-to-nonexistent relationship. The strongest relationship between OTES performance and a sub-score of teacher job satisfaction was found with Sub-score three, Security. The relationship was also considered to be weak. Additionally, Sub-score 4 was found to have a weak-to-nonexistent relationship, as well.

An analysis of the correlation coefficients for OTES Impression and Teacher Job Satisfaction proved overall weak-to-nonexistent relationships between not only overall satisfaction but for each sub-score. Overall Job Satisfaction and Supervision were found to have a direct relationship with OTES Impression. Supervision (Sub-score 1) turned out to have the strongest relationship with OTES Impression. Sub-scores 2, 3, and 4 (Responsibility, Security, and Recognition) were identified as having a weak-to-no relationship with OTES Impression. Therefore, the researcher concluded that the OTES was not associated with an impact, positive nor negative, on teacher job satisfaction.

Table 1
OTES Performance, Impression, and Job Satisfaction

	Overall satisfaction	Supervision	Responsibility	Security	Recognition
OTES Performance	0.01	-0.1	0.01	0.15	0.08
OTES Impression	0.14	0.21	0.12	-0.07	-0.15

The researcher found no statistically significant relationship between OTES performance and teacher job satisfaction. Furthermore, no statistical significance was found between any of the teacher job satisfaction sub-scores and OTES performance. Because no statistical significance was found, the researcher does not need to address the follow-up question, which asked if the relationship between the OTES and teacher job satisfaction was positive or negative. In response to the first research question, "Is the OTES associated with an impact on teacher job satisfaction?" the data showed no significant relationship between the two variables.

In order to further understand the relationship between teacher job satisfaction and the OTES, the researcher asked, "Which components of the OTES, if any, are most associated with teacher job satisfaction?" As with research question 1, a correlation matrix was created.

Table 2
OTES Components and Job Satisfaction

	Overall satisfaction	Supervision	Responsibility	Security	Recognition
Student Growth	0.03	0	-0.14	-0.06	-0.07
Pre-Conference	-0.13	0	0.02	-0.08	-0.22
Post Conference	0.19	0	-0.01	-0.10	-0.22

Overall teacher job satisfaction and each of the four sub-scores were included in the matrix. However, because overall satisfaction scores and sub-scores were already determined, additional calculations were not required. Initially, the researcher examined the student growth component of the OTES and its impact on teacher job satisfaction. Next, the researcher looked at the pre-conference and its influence on teacher job satisfaction. Finally, the researcher looked at the post-conference and if it affected teacher job satisfaction.

With regard to research question 2, "Which components of the OTES, if any, are most associated with teacher job satisfaction?" the data did not expose an association, either direct nor indirect. Therefore, the researcher concludes the OTES has had no impact on teacher job satisfaction, nor were the OTES components associated with an impact on teacher job satisfaction. In order to allow for greater insight into the impressions of the teachers who participated in the survey, participants were provided with the opportunity to make open-ended comments. Interestingly, while most open-ended comments regarding the OTES were negative; according to the data, the negative impressions and feelings conveyed in the open-ended comment box did not impact teacher job satisfaction. As with research question 1, the correlations for research question 2 revealed no statistically significant relationships.

Discussion

Legislators, superintendents, and principals may take satisfaction in knowing that by implementing the OTES, teacher job satisfaction has not been negatively impacted. However, if the desired outcome is to improve teacher job satisfaction, educational leaders might pursue recommendations based on research. Based on the conclusions of the current research, the following recommendations are proposed:

1. Make the student growth measure component homogenous for all teachers. Until all teachers have value-added data opportunities, have all teachers write SLOs to show student growth. This would alleviate the claim that the OTES is unfair for some teachers and provides an advantage for others.
2. Seek ways to improve the pre-conference component of the OTES for a more meaningful and beneficial experience for teachers. If the pre-conference is a more successful tool for improving instruction, it might have a significant and positive impact on teacher job satisfaction.
3. Seek methods for improving the post-conference as a tool for improving instruction and make it a more meaningful experience for teachers. By doing so, the post-conference may significantly and positively impact teacher job satisfaction.

As the researcher sought answers to her questions regarding the OTES and teacher job satisfaction, many more questions emerged. Indeed, a multitude of future research opportunities connected to the current study exist. These questions were framed as opportunities for future research.

The OTES was initially implemented as a pilot in the 2011-2012 school year (Ohio Department of Education, 2012). Due to collective bargaining agreements that were already in place, many districts did not implement the OTES until 2013-2014 (Ohio Department of Education, 2011). Because the OTES was relatively new at the time the study was conducted and humans, in general, often resist change, the researcher suggests replicating the study at a later date. The purpose for conducting identical research at a later date would be to determine if the element of change due to OTES being new influenced responses to the survey, and therefore results, or if the actual design of the

evaluation system influences teacher responses to the survey. With these questions in mind, replicating the current research would be prudent.

The researcher used the TJSQ and researcher-scripted questions to survey the participants. Another tool that might have been used to provide insight was the Teacher Evaluation Profile (TEP). The TEP is a tool for collecting and recording data (Stiggins & Nickel, 1988). According to Stiggins and Nickel (1988) conditions must be conducive in order for growth to be possible. The TEP produces information regarding the environment for teacher evaluation thus allowing for those who use the questionnaire to examine the potential for growth within the evaluation system (Stiggins & Nickel, 1988). Understanding the teacher evaluation environment would provide a different perspective than that provided by the TJSQ. Therefore, replicating the current research with the added element of the TEP is a future research opportunity.

Another avenue to be explored is the relationship between various demographic data with teacher job satisfaction and/or the OTES. A plethora of demographic data was collected during the research process but was not used for the purpose of the current study. Especially enlightening to teacher perceptions would be analyzing the socioeconomic description of the area. Additionally, evaluating the data to see if the teacher was designated as Type A, Type B or Type C would be informative. Teachers commented on the lack of equity in evaluation due to these designations in the open-ended comments section of the survey used in the current study. Undoubtedly, the demographic data provide ample opportunities for future research.

Because the current study was a convenience sample, it might be simulated on a larger scale using a statewide database of teachers in order to determine if the regional nature of the surveyed population impacts the results. Teachers in the northwest region of Ohio were targeted because a convenience sample was used for the quantitative study. Superintendents and principals were sent an e-mail with the survey attached. The e-mail requested they forward the survey to their teachers. Unfortunately, accessing all teachers' email addresses in the state of Ohio would be time-prohibitive. Without a doubt, replicating the current study on a statewide basis would add to the value of the current research.

For the purposes of the current study, the researcher chose to survey participants on the student growth measures, pre-conference, and post-conference components of the OTES. Future research could ask participants in the survey to respond to questions regarding other components of the OTES, such as walk-through evaluations, formal observations, and growth/improvement plans. Additionally, teachers whose district employs the original 50/50 structure of the OTES might be compared to teachers whose districts use the alternative structure of the OTES. Such a comparative study would provide further information and insight into the OTES and its impact on teacher job satisfaction. Regardless, careful consideration of teacher evaluation and its impact on teacher job satisfaction is suggested. Through careful analysis of teacher feedback and appropriate action taken based on the feedback, teacher evaluation may be useful in increasing teacher job satisfaction.

References

Achieve, Inc. (2007). *Creating a world-class education system in Ohio*. Washington, DC: Achieve, Inc. Retrieved from http://www.achieve.org

Boyd, R. T. C. (1989). Improving teacher evaluation. *Practical Assessment, Research and Evaluation, 1*(7).

Borman, G. D., & Kimball, S. M. (2005). Teacher quality and educational equity: Do teachers with higher standards-based evaluation ratings close student achievement gaps? *The Elementary School Journal, 106*(1), 3-20.

Cantrell, S., & Kane, T. J. (2013). *Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET project's three-year study* (MET Project Research Paper). Seattle, WA: Bill & Melinda Gates Foundation. Retrieved from http://www.metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief. pdf.

Coker, H., Medley, D. M., & Soar, R. S. (1980). How valid are expert opinions about effective teaching? *Phi Delta Kappan, 62*(2), 131-134.

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

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

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

Evans, L. (2000). The effects of educational change on morale, job satisfaction and motivation. *Journal of Educational Change, 1*, 173-192.

Everson, K. C., Feinauer, E., & Sudweeks, R. R. (2013). Rethinking teacher evaluation: A conversation about statistical inferences and value-added models. *Harvard Educational Review, 83*(2), 349-370.

Franco, S., Zigler, T., & Lindsey, J. (2013). *Impact of the relationship between OTES and OPES on teacher and principal evaluations. (#PB-2013-08)*. Columbus, OH: Ohio Education Research Center. Retrieved from https://oerc.osu.edu/sites/oerc.osu.edu/files/research/in-progress/OERC_ResearchBrief_OTESOPESRel_6.13.13_Final_PB-2013-08.pdf

Fuming, X., & Jiliang, S. (2007). Research on job satisfaction of elementary and high school teachers and strategies to increase job satisfaction. *Chinese Education and Society, 40*(5), 86-96. doi:10.2753/CED 1061-1932400509

Guilfoyle, C. (2006). NCLB: Is there life beyond testing? *Educational Leadership, 64*(3), 8-13. Retrieved from http://www.csun.edu/~krowlands/Content/SED610/NCLB/NCLB%20from%202007/life%20beyond%20testing.pdf

Hall, A. T., Zinko, R., Perryman, A. A., & Ferris, G. R. (2009). Organizational citizenship behavior and reputation: Mediators in the relationship between accountability and job performance and satisfaction. *Journal of Leadership & Organizational Studies, 15*(4), 381-392. doi:10.1177/1548051809331504

Harris, G. (2015). *Ohio's teacher evaluation system (OTES): Early results*. The Fordham Institute. Retrieved from http://edexcellence.net/articles/ohio%E2%80%99s-teacher-evaluation-system-otes-early-results

Hochwarter, W., Ferris, G., Gavin, M., Perrewe, P., Hall, A., & Frink, D. (2007). Political skill as a neutralizer of felt accountability—job tension effects on job performance ratings: A longitudinal investigation. *Organizational Behavior and Human Decision Processes, 102*(2007), 226-239.

Iwanicki, E. F. (2001). Focusing teacher evaluation on student learning. *Educational Leadership, 58*(5), 57-59.

Johnson, S. M. (1990). *Teachers at work: Achieving success in our schools*. Scranton, PA: Harper Collins Publishers.

Kane, T. J., & Staiger, D. O. (2012). Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains. (Research Paper). MET Project. Bill & Melinda Gates Foundation.

Kellaghan, T., Stufflebeam, D. L., Pearlman, M. & Tannenbaum, R. (2003). Teacher evaluation practices in the accountability era. *International Handbook of Educational Evaluation, 9*, 609-641.

Kupermintz, H. (2003). Teacher effects and teacher effectiveness: A validity investigation of the Tennessee value added assessment system. *Educational Evaluation and Policy Analysis, 25*(3), 287-298.

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin, 125*(2), 255-275.

Lester, P. E. (1982). *Teacher job satisfaction questionnaire*. Long Island University, Brookville, NY.

Levit, M. (1972). The ideology of accountability in schooling. *Educational Studies, 3*(3), 133-140.

Loup, K., Garland, J., Ellett, C., & Rugutt, J. (1996). Ten years later: Findings from a replication of a study of teacher evaluation practices in our 100 largest school districts. *Journal of Personnel Evaluation in Education, 10*(3), 203-226.

Lynn, A. R. (2013). Teacher evaluations based on student testing: Missing an opportunity for true education reform. *Texas Journal on Civil Liberties & Civil Rights, 18*(2), 204-234.

Marzano, R. J., Frontier, T., Livingston, D. (2011). *Effective supervision: Supporting the art and science of teaching*. Alexandria, VA: ASCD.

Maslow, A. H. (1943). A theory of human motivation. *Psychological review, 50*(4), 370.

McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003). *Evaluating value-added models for teacher accountability*. Santa Monica: Rand Corporation. Retrieved from http://www.rand.org/pub/monographs/2004/RAND_MG158pdf

Medley, D. M., & Coker, H. (1987). The accuracy of principal's judgments of teacher performance. *Journal of Educational Research, 80*(4), 242-247

OCTELA (2012, July). *Summary of HB 153, SB 316 and Pension Bills Requirements with Effective Dates.* Retrieved from: http://www.octela.org/_resources/Summary%20of%20HB%20153%20and%20SB%20316%20Legislation%20and%20Effective%20Dat.pdf

Ohio Department of Education. (2007). *Ohio Standards for Educators*. Retrieved from https://education.ohio.gov/getattachment/Topics/Teaching/Professional-Development/Master-Teacher/Teacher_Standards_Quick_View.pdf.aspx:

Ohio Department of Education. (2011). *Frequently asked questions about HB153 and evaluation based compensation and seniority*. Retrieved from https://education.ohio.gov/getattachment/Topics/Teaching/Educator-Evaluation-System/District-Educator-Evaluation-Systems/Final_FAQ_for_HB153_OTES_OPES_Evaluations_Performance_Based_Compensation_and_Seniority-links.pdf.aspx

Ohio Department of Education. (2012). *Teacher evaluation system (OTES) pilot.* Retrieved from https://education.ohio.gov/getattachment/Topics/Teaching/Educator-Evaluation-System/Ohio-s-Teacher-Evaluation-System/Additional-Information/5150-OTES-Final-Report-Appendices.pdf.aspx

Ohio Department of Education. (2013). *Educator evaluation overview*. Retrieved from http://education.ohio.gov/Topics/Teaching/Educator-Evaluation-System/Educator-Evaluation-Overview

Ohio Department of Education. (2014). *Changes to Ohio Teacher Evaluation System for 2014-2015*. Retrieved from http://education.ohio.gov/Topics/Teaching/News/Changes-to-Ohio-Teacher-Evaluation-System-for-2014

Ohio Department of Education. (2015a). *Alternative Framework.* Retrieved from https://education.ohio.gov/getattachment/Topics/Teaching/Educator-Evaluation-System/Ohio-s-Teacher-Evaluation-System/OTES-Alternative-Framework-Graphic112015.pdf.aspx

Ohio Department of Education. (2015b). *OTES Model Packet.* Retrieved from http://education.ohio.gov/getattachment/Topics/Teaching/Educator-Evaluation-System/Ohio-s-Teacher-Evaluation-System/Teacher-Performance-Ratings/OTES-Model-122315.pdf.aspx

Ohio Department of Education. (2016). *State Test Updates for 2015-2016.* Retrieved from: http://education.ohio.gov/Topics/Testing/State-Test-Updates-for-2015_2016

Persevica, A. (2011). The significance of the teachers' job satisfaction in the process of assuring quality education. *Problems of Education in the 21st Century*, *34*, 98-109.

Peterson, K. D. (2000). *Teacher evaluation: A comprehensive guide to new directions and practices.* Thousand Oaks, CA: Corwin Press.

Peterson, K. (2004). Research on school teacher evaluation. *NASSP Bulletin, 88*(639), 60-79. doi: 10.1177/019263650408863906

Popham, W. J. (2000). Big change questions: "Should large-scale assessments be used for accountability?" Answer: Depends on the assessment, silly! *Journal of Educational Change*, *1*, 283-289.

Ramirez, A. (1999). Assessment-driven reform: The emperor still has no clothes. *Phi Delta Kappan*, *81*(3), 204. Retrieved from http://metis.findlay.edu:2150/stable/20439621?seq=1&

Scott, C., Stone, B., & Dinham, S. (2001). "I love teaching but…." International patterns of discontent. *Education Policy Analysis Archives, 9*(28), 1-18.
Scriven, M. (1981). Summative teacher evaluation. In J. Millman (Ed.), *Handbook of teacher evaluation* (pp. 244-271). Beverly Hills, CA: Sage.
Seashore Louis, K., Febey, K., & Schroeder, R. (2005). State-mandated accountability in high schools: Teachers' interpretations of a new era. *Educational Evaluation and Policy Analysis, 27*(2), 177-204.
Stenning, P. C. (Ed.). (1995). *Accountability for criminal justice*. Toronto, Canada: University of Toronto Press.
Stiggins, R. J., & Nickel, P. (1988). The teacher evaluation profile: A technical analysis. *Journal of Personnel Evaluation in Education, 2*(2), 151-165.
Stodolsky, S. S. (1984). Teacher evaluation: The limits of looking. *Educational Researcher, 13*(9), 11-18.
Stronge, J. H., & Tucker, P. D. (2000). *Teacher evaluation and student achievement*. Student Assessment Series. Annapolis Junction, MD: NEA Professional Library.
Tetlock, P.E. (1992). The impact of accountability on judgment and choice: Toward a social contingency model. In M. P. Zanna (Ed.) *Advances in Experimental Social Psychology* (Vol. 25, pp.331-376).
Urdan, T. C., & Paris, S. G. (1994). Teachers' perceptions of standardized achievement tests. *Educational Policy, 8*(2), 137-156.
Van den Berg, R. (2002). Teachers' meanings regarding educational practice. *Review of Educational Research, 72*(4), 577-625.
Weiss, E. M., & Weiss, S. G. (1998). *New directions in teacher evaluation.* Eric Digest, Washington, DC Clearinghouse on Teaching and Teacher Education. Retrieved from http://www.ied.edu.hk/edchange/resource/education4_2_8.html
Wood, A., & Winston, B. E. (2005). Toward a new understanding of leader accountability: Defining a critical construct. *Journal of Leadership an Organizational Studies, 11*(3), 83-94.
Zembylas, M., & Papanastasiou, C. (2005). Modeling teacher empowerment: The role of job satisfaction. *Educational Research and Evaluation, 11*(5), 433-459. Doi: 10.1080/13803610500146152