TEACHERS’ PERCEPTIONS OF THE NEW MASSACHUSETTS TEACHER EVALUATION INSTRUMENT AND PROCCESS ON INSTRUCTIONAL PRACTICE

A thesis presented
by
David L. Thomson
to
The School of Education
In partial fulfillment of the requirements for the degree of Doctor of Education
in the field of Educational Leadership

College of Professional Studies
Northeastern University
Boston, Massachusetts
November 2013
Abstract

This mixed methods study explored whether teachers who have experienced the new Massachusetts evaluation system as a member of three pilot or early adopter districts perceive it as a valuable process. The Race to the Top federal grant process required states to redesign evaluation systems that fostered effective teaching and included student achievement data in educator evaluation systems. Massachusetts designed a model system that incorporated teacher self-reflection based on a comprehensive teacher evaluation rubric, the setting of SMART goals to improve professional practice and student achievement, and evaluations based on the comprehensive teacher evaluation rubric which represents the standards and indicators of effective teaching. In the future, measures of student achievement and student surveys will become part of the evaluation system. Teacher perceptions were measured using survey and focus group data to describe their experiences with the pilot of the new Massachusetts evaluation system. Research questions included the following: how do teachers perceive the new Massachusetts teacher evaluation system as a valid and useful tool for changing their teaching practice?; how do teachers perceive the evaluation process of self-reflection, goal setting, and evaluation using the rubric as valid and worthwhile assessment of effective teaching?; and how do teachers perceive their use of the tool and engagement in the process as an effective means toward changing their teaching practice for greater student achievement? Results overall were mixed with some groups reporting positively and others less so. There were significant differences found between two districts and between the high school and middle school categories. It is unclear if these differences were due to district level experiences or differences between the levels of school. Understanding the differences in teacher perceptions and its ramifications will foster future positive implementations and improvements in teacher
performance and, perhaps, student achievement. Positive improvements in teaching practice hold the potential to increase teacher quality and student achievement. If educators do not see the new evaluation system as improving practice, then the system should be adjusted to ensure this essential aspect is achieved.

Key words: Massachusetts model evaluation system, SMART goals, comprehensive teacher evaluation rubric, student achievement
Acknowledgements

It is with deep thanks and appreciation that I acknowledge my advisor, Dr. Chris Unger. His consistent support, insight, perspective, and positive encouragement were invaluable in this process, and I know without his influence I would not have completed this journey. I would also like to thank Dr. Sara Ewell, and Dr. Chris Martes, my second and third readers, for their time, effort, and input in finalizing my thesis.

I would also like to thank my family, especially my wife Melissa and my two daughters, Madison and Molly. I appreciate your understanding, love, and support without which I do not know how I would have completed this journey. To my Dad and late Mom, Lee and Sue Thomson, I want to thank you for teaching me the value of hard work and consistent effort and for raising my brother and I to never be complacent and to continually work to improve ourselves. I also would like to thank my mother in law, Judy Deal for all of her encouragement and support over the past three years. I would also be remiss if I did not thank my friend Kathy Cook for all of her suggestions and for helping me to become a better writer.

I would also like to thank all the administrators in the early adopter districts who assisted me in gaining access to these districts, and beating the bushes to collect this data. I would be remiss if I did not mention the greatest assistant principal in the world, Richie Florence, for his never-ending understanding, support, and encouragement.
Table of Contents

Abstract ........................................................................................................................................ 2
Acknowledgements ...................................................................................................................... 4
List of Tables .................................................................................................................................. 8
Chapter I: Introduction .................................................................................................................. 9
  Problem of Practice and Significance ......................................................................................... 9
  Significance ................................................................................................................................. 10
  Practical and Intellectual Goals ................................................................................................. 11
  Research Questions .................................................................................................................... 13
  Theoretical Framework ............................................................................................................... 13
Chapter II: Literature Review ...................................................................................................... 16
  The Rationale for Implementing a New State-wide Teacher Evaluation System ...................... 16
  Definitions of Effective Teaching and How Effective Teaching Should Be Measured ............... 18
  Definitions and Examples of Effective Evaluation Models ....................................................... 20
  Characteristics of the New Massachusetts Educator Evaluation Model .................................... 21
  Research on the Impact of Evaluation Models on Teachers’ Instructional Practice ................. 22
  Research on the Relationship Between Educator Evaluation and Student Performance .......... 23
  Summary ..................................................................................................................................... 25
Chapter III: Methodology ........................................................................................................... 26
  Research Questions .................................................................................................................... 26
  Qualitative Methodology .......................................................................................................... 27
  Site and Participants .................................................................................................................... 28
  Data Collection .......................................................................................................................... 29
  Data Analysis ............................................................................................................................... 30
  Validity and Credibility ............................................................................................................... 30
  Protection of Human Subjects and Ethical Considerations ....................................................... 31
  Conclusion .................................................................................................................................. 32
Chapter IV: Research Findings ................................................................................................... 32
  Reporting of the Findings and Analysis ...................................................................................... 32
  Study Context ............................................................................................................................... 33
  Survey Participants ..................................................................................................................... 33
  Survey Analysis ........................................................................................................................... 35
  Focus Groups and Interview Participants .................................................................................. 35
  Focus Group and Interview Data ............................................................................................... 35
Research Question #1: Do Teachers Perceive the New Massachusetts Teacher Evaluation System as a Valid and Useful Tool for Changing their Teaching Practice? .................................................. 36

Did SMART goals, as part of the new teacher evaluation process, help teachers focus on the improvement of their teaching practice? ................................................................. 36

The new evaluation process resulted in changes in my teaching methods ................................................. 39

My evaluator was able to gather sufficient evidence to accurately assess my performance .............................. 42

Looking back, the evaluation process resulted in positive changes in my professional practice.......................... 45

Research Question #2: How do Teachers Perceive the Evaluation Process of Self-reflection, Goal setting, and Evaluation using the Rubric as Valid and Worthwhile Assessment of Effective Teaching? .................................................. 49

The self-reflection process caused serious consideration of individual’s teaching practice .............................. 49

The rubric, the standards and indicators of effective teaching practice, is a comprehensive guide of the traits of effective teaching.................................................. 51

Overall, what are your thoughts on the new teacher evaluation system? ................................................ 54

What do you like about the new teacher evaluation system? ................................................ 55

Research Question #3: Do teachers perceive use of the tool and engagement in the process as an effective means toward changing their teaching practice for greater student achievement? ..................... 57

Summary of Findings .......................................................................................................................... 60

Chapter V: Discussion of the Research Findings .................................................................................. 63

Summary of the Problem of Practice .................................................................................................. 64

Review of the Methodology ................................................................................................................. 64

Summary of the Findings ...................................................................................................................... 64

Discussion of the Findings in Relation to the Theoretical Framework ................................................. 64

Discussion of the Findings in Relation to the Literature Review ...................................................... 73

Definitions of effective teaching and how effective teaching should be measured ......................... 74

Definitions and examples of effective evaluation models ................................................................. 74

The impact of evaluation models on teachers’ instructional practice .............................................. 75

Validity and Limitations ...................................................................................................................... 76

Significance of the Study in the Field .................................................................................................. 77

Final Words ........................................................................................................................................... 78

Next Steps ............................................................................................................................................ 80

References ............................................................................................................................................. 83

Appendix A ........................................................................................................................................... 90

Appendix B ........................................................................................................................................... 91

Appendix C ........................................................................................................................................... 92
List of Tables

Table 1  Demographics of survey participants ................................................................. 34
Table 2  SMART goals helped focus teachers on improving their teaching practice .......... 37
Table 3  The new evaluation process changed my teaching methods ............................... 40
Table 4  Evaluators were able to gather sufficient evidence to assess accurately teacher
performance. .................................................................................................................. 43
Table 5  The evaluation process results in positive changes in professional practice. ........ 46
Table 6  The self-reflection process caused serious consideration of individual’s teaching
practice. ......................................................................................................................... 50
Table 7  The rubric is a comprehensive representation of effective teaching practice ........ 52
Table 8  Setting SMART goals has help to focus my efforts on increasing student achievement
........................................................................................................................................ 58
Chapter I: Introduction

As the faculty meeting concluded - a meeting that focused on the new Massachusetts evaluation process, the principal was summoned to an impromptu meeting with teachers representing multiple disciplines. The anxiety and distress were evident as the teachers shared their concerns regarding their ability to meet the standards set in the evaluation instrument (Department of Elementary and Secondary Education, 2011). These concerns, replicated throughout the Commonwealth, are due to factors such as class size, special education students, and limited resources available to educators. Conversely, existing evaluation tools as well as how the instruments are used to evaluate teacher performance, do little to improve teacher instruction. A study of 309 school districts conducted by Weisberg et, al., 2009, reported the following: nearly all teachers received good or excellent ratings; districts failed to recognize excellent teaching; professional development was rarely tied to teacher evaluation results; and poor evaluations rarely led to teacher dismissal. While it is much too early for measurable results in regards to improving student achievement, it is important to examine the role of teacher attitudes and perceptions in this process. Hall (2008) found that when examining teachers’ response to high stakes evaluation systems, if teachers believe that the goals are valid and the process effective, they will work toward the improvement goals.

Problem of Practice and Significance

Research has demonstrated that the most important factor in student achievement is the quality of the classroom teacher (Marzano, 2003; Hanushek 2009; Rothstein, 2004). Quality of instruction directly influences student achievement (Ripley, 2010). Although there is little disagreement of the importance of the teacher, there are different opinions about teacher effectiveness: how to define it, as well as how to measure it.
The state of Massachusetts is now in the third year of implementing a new teacher evaluation model. This instrument identifies qualities of effective teaching based on Danielson’s framework for teaching (2007). The instrument evaluates teachers by rating various components of instruction using a comprehensive rubric. This rubric characterizes the qualities of an effective teacher in four areas: 1. curriculum, planning, & assessment; 2. teaching all students; 3. family & community engagement; and 4. professional culture (Department of Elementary and Secondary Education, 2011).

The goal of the impending evaluation process is to provide a focus for the improvement of teaching practices and ensure that only highly effective individuals lead instruction in our classrooms (Rothstein, 2010; Massachusetts Department of Elementary and Secondary Education, 2011). The Race to the Top grant process has mandated educator evaluation systems that foster effective teaching and incorporate student achievement scores (U.S. Department of Education, 2009). Many researchers, including Darling-Hammond, have advocated for evaluating the characteristics of effective teaching based on research that established characteristics and methods of quality instruction (1999). Danielson promotes the use of a common framework to guide professional conversations surrounding effective teaching (2007). The new Massachusetts model for educator evaluation represents a significant shift in the evaluation process. The rubrics and other documents that have been created by the state will not only serve as evaluation instruments but are also intended to serve as a structure and basis for educators to self-assess their practice and reflect on the effectiveness of their instruction. (Massachusetts Department of Elementary and Secondary Education, 2012).

**Significance** The new evaluation process will have the potential to be used in high stakes personnel decisions, eventually based in part on data culled from student assessments.
The question is whether the teachers perceive the new evaluation process as resulting in changes in their teaching practice that eventually will lead to higher student achievement. Cognizant that the Massachusetts model is being implemented in phases, investigating the early stages of this implementation – specifically teacher perceptions, their motivation and investment in the process, as well as the organizations’ clarity of the goals in the evaluation process – can inform the more effective implementation of the evaluation process going forward. In many districts, the new evaluation model is a drastic departure from previous forms of evaluation. One aspect that is revolutionary is the use of a rubric of effective teaching characteristics as well as goal setting with measurable outcomes geared towards increasing student achievement and improving professional practice. The implementation of the new system is on a fast tract, leaving little time for training for both teachers and evaluators. If teachers are not invested in the process or do not demonstrate a belief in the goals or feedback of the system, then performance improvement of instruction will be little or nonexistent. This could negatively affect the implementation and ultimately the overall success of the initiative. Furthermore if behavior is not changed by the new system, then instructional improvement and thus increased student achievement will not occur. This study examines the teachers’ perceptions of the new evaluation system and indicates the capacity of the system to improve teacher practice.

**Practical and Intellectual Goals**

This study examined the perceptions of teachers during the initial implementation of the Massachusetts evaluation model. The results provided information on the perceived validity of the instruments and evaluation process and if, in the teachers view, the implementation of the instruments caused changes and improvements to their teaching practices. Measuring teacher effectiveness is a complex task. Developing effective teaching practices begins with the
understanding of these practices as well as the importance of implementing these essential skills.

Given the problem of practice, the most essential goal was to see if the evaluation process, which incorporates educator self-reflection and goal setting, in fact resulted in changes to teaching practice and fostered effective teaching. While that question cannot be asked at the time of this research, one can inquire as to the degree that early users of the instrument and process perceived the instrument and process as in fact effectively increasing their self-reflection and teaching practice. This led to the following intellectual goals:

1. To assess the degree to which early users of the new Massachusetts evaluation model perceived the process as a vehicle for effective self-assessment, reflection, and change in their teaching practice.

2. To assess whether the use of the observation data, teacher provided evidence, and goal setting, in its present form in the Massachusetts evaluation process, was perceived by educators using the system as reliable in the assessment of educator quality.

The practical goal of the study was then to determine whether there were aspects of the current evaluation process that teachers felt did and/or did not effectively support their reflection on teaching, goal setting, and instructional practice. Assessing teachers’ perspectives on these intended outcomes of the process was expected to yield helpful information as to how the process or teachers’ engagement in the process could or should be revised or refined to be more effective toward its intended outcomes.

Through this research, one will examine the teachers’ perceptions of the Massachusetts evaluation model to determine if teachers viewed the evaluation model as a valid and effective
method of evaluation, and in response to the new model, reported changing their teaching practice.

**Research Questions**

This qualitative study examined the perceptions and opinions of educators implementing the new evaluation model in Massachusetts in the first year of implementation. This information may inform the future success of this process as it is implemented throughout the Commonwealth.

The central questions in this research are:

1. Do teachers perceive the new Massachusetts teacher evaluation system as a valid and useful tool for changing their teaching practice?
2. How do teachers perceive the evaluation process of self-reflection, goal setting, and evaluation using the rubric as valid and worthwhile assessment of effective teaching?
3. Do teachers perceive their use of the tool and engagement in the process as an effective means toward changing their teaching practice for greater student achievement?

**Theoretical Framework**

In order to build further understanding of the many factors influencing the new systems of educator evaluations in Massachusetts, this study was approached from the framework of Performance Improvement Theory (Swanson, 1999). In Performance Improvement Theory (PI), one is focused on clarifying the goals and behaviors, as well as harmonizing the goals and behaviors of individual contributors, work process owners, and organization leaders (Swanson, 1999).
The Massachusetts evaluation system has identified characteristics and behaviors that are essential to effective teaching: the new evaluation system identifies these traits in the teacher performance rubric. In order for the new evaluation tool to be effective, educators will need to invest themselves in the process and become cognizant and aware of the need to change their practice in reference to the evaluation process, which includes both goal setting and the described characteristics of effective teaching. The Massachusetts evaluation system process requires evaluators to establish goals and discuss progress during the evaluation cycle. As Swanson describes, performance improvement is dependent on goal clarity and the changing of behaviors of those asked to improve their performance (1999). Smither, London, and Reilly state that personality and the orientation of feedback can greatly affect the perceived need to change one’s performance (2005). To be more specific, Smither et al. found that feedback alone is not the cause of behavior change, but that receiving negative or constructive feedback and developing goals from that type of feedback institute changes in performance (2005). Schaffer and Keller found that stakeholders should be involved in the design of performance evaluation systems early to ensure success (2003). Furthermore, Schaffer and Keller contend that an evaluation process should incorporate existing measurement processes instead of creating new systems, ensuring continuity of the program (2003). Smither et al. suggest that using feedback to guide the recipient’s development and learning goals will enhance the subsequent performance improvement of the individual (2005). These finding are consistent with the mid-year formative assessment process in the Massachusetts evaluation model (Massachusetts Department of Elementary and Secondary Education, 2012).

Swanson suggests that the human capital is most important in this process, and that performance cannot be improved if people choose not to perform (1999). This can be attributed
to a systematic disconnect in the form of ambiguous outcomes and the lack of a systematic performance improvement process (Swanson, 1999). Smither et al. found that when the purpose of feedback was developmental and not administrative in nature, there were less bias and fewer rating errors, providing increased accuracy in evaluations, and greater performance gains (2005). Furthermore, the performance improvement framework suggests leaders need to clarify the goals and develop the behaviors of individuals in the system (Swanson, 1999). It is important to note that performance improvement is probable for those individuals who take action based on the feedback received (Smither et al., 2005). The importance of support, made available in the form of coaching, group discussions, and developmental activities, ensures appropriate action will be taken to improve performance (Smither et al., 2005). In order to experience changes and improvement in instructional practice of teachers, one would expect to see evidence of the implementation of this Performance Improvement framework in the Massachusetts evaluation model.

Fundamentally, there is a universal effort to increase the academic performance of the nation, especially the underserved populations, through increasing teacher effectiveness and capacity (A Nation at Risk, 1983: Race To The Top, 2009). Previous attempts, specifically the use of merit pay, have been made to improve education and have not demonstrated significant gains in student performance (Burns, Gardner, & Meeuwse, 2009; Glazerman & Seifullah, 2010; Springer, Ballou, Hamilton, Le, Lockwood, McCaffrey, Pepper, & Stecher, 2010). Through the lens of performance improvement theory, this study focused on the teacher perceptions of the clarification of goals and behaviors, as well as the harmonizing of these goals. The behaviors of teachers and principals as well as the teachers’ investment and motivation in
the new evaluation process will indicate the likelihood of the overall success of the evaluation model.

**Chapter II: Literature Review**

Based on this proposal and the research questions the literature had been reviewed in the following areas: the rationale for implementing new teacher evaluation systems, effective teaching, the use of effective evaluation models, the Massachusetts evaluation system, and linking teacher evaluations to student performance. The intent of the review was to provide context and content to the purpose, design, and analysis of this study in relationship to previous and current scholarly literature and research in this area.

**The Rationale for Implementing a New State-wide Teacher Evaluation System**

The Race to the Top grants have brought a revolutionary and controversial concept to the forefront of the field of education; the measurement of educator effectiveness based, at least in part, on student achievement data. Each district and state receiving these federal funds must revise their teacher and principal evaluation systems to include student achievement data. This emphasis was then included in the No Child Left Behind (NCLB) waivers granted by the United States Department of Education (2012). These efforts are part of the response to the concern that a decrease in educational quality relative to other countries will affect the economic growth in our country (Hanushek 2009). This echoes the call to arms raised in 1983 in *A Nation at Risk*, a landmark report released by a panel assembled by President Reagan. The rationale for pursuing the measurement of student outcomes and using this information to evaluate educators is research that indicates that the single most important factor in student achievement is the quality of the teacher in the classroom (Marrzano 2003, Hanushek 2009). In fact, New Leaders for New Schools stated that the impact of an effective teacher is two and a half times greater than that of
small class size (2009). Furthermore, in How to Improve the Supply of High-Quality Teachers edited by Rothstein (2004), Hanusek and Rivkin make the case that a top quality teacher can elicit one and a half grade levels of growth in a single year. Proponents argue that by measuring what the students learn one can judge the effectiveness of teachers and their instruction. This information could empower school districts to remediate ineffective teachers or remove them from the classroom (Rothstein, 2010, Massachusetts Department of Elementary and secondary Education, 2011).

Our present system of evaluation is essentially a formality based on prescheduled observations conducted by administrators and is often not always reflective of teaching and learning (Holland, 2006). More often than not, the supervision model includes pre- and post-observation meetings with a scheduled observation. These systems of evaluation are often not tied or aligned to teaching standards or student outcomes (O’Day, 2002).

Chetty, Friedman, and Rockoff (2012) found that value added measures accurately predict student achievement on test scores when student characteristics are controlled. Value added is the amount students’ gain in their test scores and is attributed to the effect of the teacher on student achievement. Chetty and his colleagues found that teachers with high value added scores have students that are more apt to attend college, earn more, live in better neighborhoods, save more for retirement, and are less likely to become teenage parents. In contrast to criticism regarding value added measures (Newton, Darling-Hammond, Haertel, & Thomas, 2010: Ewing, 2011), Chetty and his colleagues state that value added measures are not biased and in fact separate the teachers’ effect from other influences on student achievement (2012). Chetty et al. present that higher value added teachers not only increase test results, they also improve student
outcomes in life. Improved education through improved teaching would increase the overall success of individuals as well as the state of the country as a whole.

**Definitions of Effective Teaching and How Effective Teaching Should Be Measured**

Many would agree that an effective teacher works in such a way that their students show gains in achievement measures, often documented in the form of value-added scores on state tests (Gates Foundation, 2009). However, there are many opinions as to what specific characteristics are inherent in effective teachers. Berry, Daughtrey, and Wieder (2010) proposed that soliciting parent involvement, communicating positive expectations, improved instructional practice, and willingness to innovate in the classroom are important characteristics of effective teachers. Giovannelli (2003) suggests five components of effective teaching: classroom management, instructional behavior, classroom organization, teacher expectations, and the composite of effective teaching.

Porter and Brophy (1998) suggest that effective teachers are knowledgeable about their content and strategies to teach it, are clear in their instructional goals and expectations, know their students and adapt instruction to their needs, monitor student understanding, offer regular appropriate feedback, address higher as well as lower cognitive objectives, enrich and clarify the content, accept responsibility for student outcomes, and are thoughtful and reflective about their practice. Mortimore (1994) summarized the skills of effective teachers as organizational, analytical, synthesizing, presentational, assessing, managerial, and evaluative. Organizational is the ability to sort out materials and sources of information. Analytical is the skill of breaking down complex sources of information. Synthesizing is the bringing together of ideas and content into understandable lines of reasoning. Presentational skill is the ability to clarify complex information without changing its meaning. The skill of assessing refers to the ability to judge the
work of student and give appropriate feedback. Managerial skill is the need to coordinate the
dynamics of the individuals in the classroom. An evaluative ability is the self-assessment
necessary to improve teaching on a continual basis.

Darling-Hammond (2009) has found that teacher effectiveness is based on teacher
qualities that include intelligence, verbal ability, strong content knowledge, knowledge of how to
teach others in that content area, an understanding of learners and how they learn, adaptive
ergise and the ability to respond to students’ needs, a willingness to support student learning,
a willingness to collaborate, and lastly, the inclination to continue to learn and improve their
craft. Brophy and Good (1986) found that student achievement was higher when teachers
covered more material, class time was allocated to academic activities, students experienced
consistent success in the classroom, and teachers were actively engaged in teaching the students.

Harris (1998) states that “effective teaching is linked to reflection, inquiry, and
continuous professional development and growth” (p.179). Teachers who were either finalists or
winners of the National Teacher of the Year award were included in a qualitative study by Van
Schaack and Glick (1982) to determine their common characteristics. Van Schaack and Glick
found that these exemplary teachers possessed an enthusiasm for teaching, superb
communication skills, a Socratic approach, warmth and concern for their students, self-
confidence in their abilities, a belief that all students should and can succeed, deep
professionalism, and a commitment to renew themselves and continually improve their teaching.

Effective teaching is a function of individual characteristics, teacher behavior and skill, the
willingness to improve consistently, and an understanding of desired outcomes and goals (Harris,
1994; Darling-Hammond, 2009; Berry et al., 2010; Giovannelli, 2003).
Definitions and Examples of Effective Evaluation Models

Having high quality, effective teachers is obviously important and there is a strong position advocating that measuring student outcomes should be central in identifying effective teaching. Darling-Hammond agrees with the importance of teacher quality and its impact on student learning, but promotes policies centered on fostering teachers’ development of subject area knowledge, pedagogy, experience, and collaboration as a means to increase student achievement (1999). Kane promotes that evaluation should be approached from multiple angles including student surveys, observations, and student achievement measures, ensuring accurate findings (2012). However, Kane’s work did find that achievement gains were more prevalent in Math than English, and his data was not used in high-stakes personnel decisions (2012). In fact, Kane warned that the measurement properties of student surveys, observations, or achievement could be distorted if high stakes were attached to the outcomes of the evaluation (2012). While student achievement gains provide few clues for how to improve practice, observations by a trained adult with a well-developed instrument do provide effective feedback (Kane, 2012). Taylor and Tyler state that high quality observations provide specific information on how to improve teaching practice (2012). However, Kane highlights the need for the observation instrument to capture a set of effective teaching practices, and that the role of evaluation should be to improve practice (2012). In Taylor and Tyler’s research, they found that teachers were more effective at raising student achievement in the year they were being observed than in previous years (2012). The encouraging news is that Taylor and Tyler also found that the observed teachers were even more effective in the years following the observation year (2012).

In her study of two high performing California charter schools, Morelock (2008) found that the evaluation model they implemented led to changes in teacher behavior and improvement
in student achievement scores. The focus of these evaluation systems was on improving teacher practice through a process of self-reflection, professional development, and standards of high quality teaching practice (Morelock, 2008). Although Morelock point out that the impact of performance standards on student achievement data is inconclusive, the common understanding and expectations were essential to the program (Morelock, 2008). Morelock found significant evidence that the evaluation process led to positive changes in teacher behavior and practice, enabling the school to meet the needs of the students (2008).

The Massachusetts teacher evaluation model assesses key characteristics in the areas of curriculum, planning, and assessment, teaching all students, family and community engagement, and professional culture (Massachusetts Department of Elementary and Secondary Education, 2012). These characteristics or qualities are supported by research based on teacher ratings and student achievement; however, these qualities present challenges in measurement and evidence.

**Characteristics of the New Massachusetts Educator Evaluation Model**

As mentioned above, the rubric developed to evaluate Massachusetts teachers is based heavily on the Danielson model (Massachusetts Department of Elementary and Secondary Education, 2012; Danielson 2007). The process begins with a self-evaluation based on the standards and indicators of effective practice and a conference where “smart” or measureable goals are established. One of these goals is to be focused on student achievement and another on the professional practice of the educator. Evaluators are required to make unannounced observations and use this information in combination with other evidence collected by teachers, as well as a goals progress review, to conduct midyear and year-end conferences leading to a summative evaluation (Massachusetts Department of Elementary and Secondary Education, 2012). The process leads to consistent dialogue regarding the evaluation evidence, professional
practice, and how this information can be assessed using the standards and indicator rubric. Reflecting on the evidence and goals, a final summative evaluation is made determining if the teacher is exemplary, proficient, needs improvement, or unsatisfactory (Massachusetts Department of Elementary and Secondary Education, 2012). In two years, 2014-2015, student data, including growth measurements calculated from state and district assessments will be used in conjunction with summative evaluation finding, to determine the cycle of evaluation and the development of the educators’ growth plans (Massachusetts Department of Elementary and Secondary Education, 2012; Massachusetts Department of Elementary and Secondary Education, 2011).

**Research on the Impact of Evaluation Models on Teachers’ Instructional Practice**

Many studies have examined merit—based systems that promote similar goals of improving teacher effectiveness. These initiatives have been used and studied in Austin, Texas and Chicago, Illinois (Burns, Gardner, & Meeuwsen, 2009, Glazerman & Seifullah, 2010). These merit-based pay systems combined measured student growth and professional development to reward teachers financially for reaching personal goals as well as reaching student performance goals (Burns et al., 2009, Glazerman & Seifullah, 2010). In Chicago, the authors found that even after two years there was no evidence that the Teacher Advancement Program increased student test scores (Glazerman & Seifullah 2010). In Austin, Texas, only 36 percent of teachers in the program who had received merit pay reported that they had changed their teaching practice (Burns et al. 2009, p. 2). Another recent study conducted in Nashville from 2006 to 2009 found that there was no significant difference in student achievement results between students whose teachers were in a pay for performance program and those that were not (Springer, Ballou, Hamilton, Le, Lockwood, McCaffrey, Pepper, & Stecher, 2010).
Furthermore, Matthew Springer and colleagues found in their multiyear study that there was no significant difference in a single or multiple years between the incentive and non-incentive groups (2010). Rothstein concluded that that on its own, merit pay does not initiate positive change in student test scores or significantly enhance teaching practice (Rothstein, 2004).

Adams found that southwest Missouri educators did not believe they experienced professional growth through their state’s performance evaluation process (2009). Despite clearly defined expectations and procedures, few felt they had become better teacher after being evaluated (Adams, 2009). Hall (2008) found that teachers would adjust their instruction and planning to increase test scores and improve their own evaluation scores. Hall (2008) also found that teachers would conform to an evaluation system that is soundly based. In examining the impact of the standards based teacher evaluation system in Iowa, Huckstadt found teachers reported that the standards did challenge them to improve their practice (2011). However, teachers reported only minimal changes in instructional practice as a result of participating in the standards-based evaluation system (Huckstadt, 2011). In their study evaluating the use of portfolios in teacher evaluation, Tucker, Stronge, Gareis, and Beers were disappointed to find that although teachers reported the portfolios encourages self-reflection, it did not lead to changes in teaching practice (2003). However, Tucker et al. did find that portfolios did enhance the evaluation of teachers for accountability and professional development (2003). The discrepancies in the research underpin the importance of performance improvement theory in the implementation of any effort to improve teacher practice and performance.

**Research on the Relationship Between Educator Evaluation and Student Performance**

Several studies have examined the effect of teacher evaluation ratings and other measures of teacher effectiveness and the relationship of these measures to student performance (Borman
& Kimball, 2005; Gallagher, 2004; Kimball, White, Milanowski, & Borman, 2004; Milanowski, 2004). Gallagher (2004) found that teachers’ influence, characterized as a classroom effect, increased student performance in the area of reading. However, results in the area of mathematics did not demonstrate a similar result due to curriculum alignment issues (Gallagher, 2004). Kimball, White, Milanowski, and Borman found that teacher evaluation scores influenced student achievement, and although a positive correlation was found, it was less than statistically significant (2004). Milanowski (2004) also found a correlation between teacher evaluation ratings and student achievement, and although the study only examined student data from one year it indicates a relationship is probable. Borman and Kimball found that better teaching leads to better than expected learning outcomes; however it was also noted that the higher rated teachers seemed unable to close achievement gaps (2005). Lyon determined that there was a strong relationship between schools ranked in the top ten in student achievement and the use of student achievement data in the teacher evaluation process (2009). Taylor and Tyler found that teachers who were evaluated had an average of a .05 standard deviation increase in their students’ math scores from the previous year (2012). In fact, teachers with lower scores in the beginning of the evaluation process had the largest growth in student achievement measures at the end of the evaluation cycle (Taylor & Tyler, 2012). Taylor and Tyler conclude that those that received the most critical feedback and had the most room for improvement demonstrated the greatest change in student outcomes and improvement in overall teaching practice (2012). These studies support the relationship between evaluation systems that measure teacher effectiveness and produce increased student performance.
Summary

In any consideration of a new evaluation system, it has been argued that a balanced approach consisting of the assessment of the act of teaching as well as the results of teaching should be considered (Tucker & Stronge, 2005). Darling-Hammond (2009) supported a standards-based approach to assessing teacher quality. In this approach, exemplified in the National Board for Professional Teaching Standards Certification Process, teachers participate in a process where they are assessed on performance-based standards that are significantly linked to teacher effectiveness. In standards-based evaluation programs, teachers are observed by trained observers and videotaped. They also engage in reflective dialogue on their practice and review evidence of student learning. This type of evaluation system focuses on teacher behaviors that promote student learning and guides teachers to evaluate their own performance. The Danielson model (2007) has been incorporated in the evaluation systems in many states across the country including Idaho, Oklahoma, Chicago, and New York City. This model sets criteria, through its rubrics, for effective teaching characteristics and facilitates a dialogue with a common vocabulary (Danielson, 2007). Darling-Hammond (2009) highlighted state systems in Minnesota and Arizona because they have developed systems that incorporate teacher collaboration, valid evidence of teacher effectiveness, standards-based evaluation of practice, evidence of student learning on a variety of assessments, and ongoing high quality professional learning. When a standards-based system is in place linking student performance objectives to a teacher compensation system, there is a positive correlation to student achievement gains (Darling-Hammond, 2009). The research by Gallagher, Kimball, Boraman, Milanowski, and White indicates that there is a relationship between effective teaching and increased student achievement (Borman & Kimball, 2005; Gallagher, 2004; Kimball, White, Milanowski, &
Borman, 2004; Milanowski, 2004). Although more research is needed, the standards-based evaluation model could present a comprehensive system of evaluation leading to improved student achievement.

Examining this problem through the theoretical framework of change and performance improvement theory, it is important to monitor and foster motivation and investment, clarity of goals and behaviors, and harmonizing the goals and objectives among the constituents in the new evaluation process. Darling-Hammond (2009) observed that these merit-based pay for performance efforts failed for the following reasons: “faulty evaluation systems, concerns about bias, the undermining of collaboration, and the lack of public will to continue the increased compensation” (p.16). In heeding the need for change and improvement, studying the implementation of the teacher evaluation system in Massachusetts will inform and improve the process.

**Chapter III: Methodology**

This study incorporated a mixed methods methodology that began with surveying teachers who had experienced the new Massachusetts Evaluation System in early adopter districts. These initial findings were followed with focus group discussions to further establish patterns and themes of the participants’ experience of the new system.

**Research Questions**

Three research questions guided the collection of data and its analysis.

1. How do teachers perceive of the evaluation system as a valid and useful tool for changing practice?
2. How do teachers perceive the evaluation process of self-reflection, goal setting, and evaluation using the rubric as a valid and worthwhile assessment of effective teaching?

3. Do teachers perceive their use of the tool and engagement in the process as an effective means toward changing their teaching practice for greater student achievement?

**Mixed Methods Methodology**

This mixed-methods study used survey and focus group data to examine teacher perceptions of the new teacher evaluation system (Fraenkel & Wallen, 2009). Fraenkel and Wallen (2009) state that a mixed methods, explanatory design allows the researcher to flush out the meaning behind or to refine quantitative data allowing the researcher to expand on the results of a quantitative study (p. 561). Participants in these early adopter districts will be actively involved in the second half of the evaluation process; therefore, they possessed both subjective and objective experiences (Creswell, 2009). Three districts that voluntarily participated in the pilot of the new evaluation system were targeted for both the survey and focus groups. Unfortunately, in two districts this researcher was limited by the number of schools that could be approached, and in the third district, by the very small number of actual participants involved in the pilot process. The survey asked participants to indicate on a five-point Likert scale their agreement with statements referring to the evaluation process, its ease of use, and their perception of how it impacted their teaching practice and student achievement. The survey also asked several open-ended questions regarding their likes and dislikes of the new process. These results were analyzed and reported, providing the percent of agreement and disagreement across all items across teachers, and the mean response and standard deviation across subpopulations of
the teachers who participated. These results were also analyzed to determine whether the differences across subpopulations of teachers were significant, using the Fisher Exact Test for significance.

In order to gain a deeper understanding of the survey results, focus groups were also conducted in each district asking participants to provide their opinions about the new evaluation process. These discussions were then transcribed and the transcriptions shared with each focus group member to review and check for accuracy. Once this process for accuracy and precision was completed, the transcripts were analyzed for significant statements and categorized with respect to what was learned and specific key themes that were pertinent and obvious within each question. These clusters of meaning from the survey and focus group sessions were then used to describe the participants experience in the form of response data and textural descriptions (Creswell, 2007). The mix of these data sources allowed the researcher to describe the experiences and perceptions of the participants and use these responses to clarify the results from the survey questions.

**Site and Participants**

The participants were public school teachers in K-12 education who had participated in the evaluation pilot in three early adopter school districts in central, greater Boston, and southeastern Massachusetts (Appendix A). These districts were outside of where the researcher works and had piloted the evaluation protocols in the previous school year. A sample of teachers currently in the field was sought to represent time in the profession, gender, and grade level taught. The survey was conducted using an electronic format, specifically survey monkey (Appendix D). Survey participants were then solicited from the three school districts through the online survey to participate in focus groups. Participants were identified with varying levels of
experience in the teaching profession, different subject areas, gender, as well as grade levels taught. This stratification provided a diverse sample of experiences and perspectives for this study. Additionally, this approach allowed possible comparisons between various subgroups of teachers based on similarities and differences. Informed consent was obtained from all participants. Rapport was developed by assuring participants that their statements would be reported anonymously. Institutional Review Board (IRB) approval was granted based in part on the fact that no more than minimal risk would exist for the participants whose privacy and confidentiality was maintained. While participants were not offered incentives for participating in the study, the findings of the study would be shared with participants.

Data Collection

The opportunity to participate in the survey was extended to all teachers who had piloted the evaluation system, and represented the three levels, elementary, middle, and secondary, in three early adopter districts (Appendix B). All participation was voluntary and abided with all expectations and guidelines of the Northeastern University Internal Review Board (Appendix C and F). The survey data was collected, collated, and analyzed and provided insight into educators’ feelings and perspectives of the new evaluation process. Participants for the focus groups were solicited from the survey respondents, and focus group questions were developed to gain further insight into themes revealed in the survey responses. The focus groups, one in each district, three focus groups in total, were held in a neutral school-based setting, after school, and were as informal as possible in an effort to alleviate subjects’ anxieties. In the focus group sessions, honest and open answers were solicited to questions based on teacher’s experiences on the Massachusetts teacher evaluation model.
Data Analysis

This researcher incorporated methods of organizing and analyzing Likert scale data consistent with Maxwell (2005) including highlighting significant findings and developing clusters of meaning. Descriptive statistics were incorporated to determine the mean and standard deviation enabling the researcher to develop assumptions of this initial set of data (Fraenkel & Wallen, 2009). These statistics indicated the general feelings and perceptions of the teachers and themes were further explored in the focus groups. (Appendix D and G).

Records of all focus group sessions were in the form of notes and digital audio recordings. In order to analyze the qualitative data from the focus groups and interviews, participant responses were first transcribed then analyzed for themes and clusters of answer types; the responses were then categorized and used to add insight and commentary to survey question results. The researcher examined this data to identify significant statements, clusters of meaning, and subsequently, themes reflecting the teachers’ feelings and perceptions of how the new evaluation system impacted their teaching practice, self-reflection, and its ability to evaluate effective teaching.

After the data was analyzed, significant statements were categorized with respect to what was learned and specific key themes that were pertinent and obvious within each question. These clusters of meaning from the survey and focus group sessions were then used to describe the participants’ experience in the form of response data and textural descriptions (Creswell, 2007).

Validity and Credibility

As the researcher is a middle school principal, evaluated with a similar evaluation instrument and methodology, he bracketed his own bias and preconceptions (Moustakas, 1994).
This researcher served as the state Secondary Principal Association president during the
development of the new evaluation system and directed the principals association’s response to
this initiative. Furthermore, this researcher has been an active member of advisory panels
providing feedback to the Massachusetts Department of Elementary and Secondary Education
department on the implementation of the new evaluation system. Lastly, as an active
administrator this researcher was responsible for implementing this system in his own building.
To reduce the effect this researcher could have on the responses provided, this research was
carried out outside of his home district in three early adopter districts where the researcher does
not hold any formal role. To increase validity and reliability, an identical interview format was
utilized with each focus group. Furthermore, all interviews took place in neutral settings. The
researcher conducting the interviews maintained an unbiased approach by not reacting to
statements made by participants that suggested agreement or disagreement, and avoided the
possibility of influencing further responses. Once this was completed, the researcher validated
the findings by meeting with participants one final time to ensure methodological rigor and
accuracy of the data (Maxwell, 2005).

Protection of Human Subjects and Ethical Considerations

Throughout this study ethics and the protection of human subjects were maintained,
namely respect for persons, beneficence, and justice. This researcher participated in and passed
the National Institute of Health’s (NIH) online course, “Protecting Human Research
Participants” certification number 810902 (Appendix I). To ensure respect for the individuals
participating in the study, participation was voluntary and to the extent possible, anonymous.
Participants in the focus groups were solicited in the survey and volunteered by indicating on the
survey their desire to participate in this forum. An informed consent form identifying the
purpose of the research as well as its risks and benefits was used. The researcher also guaranteed confidentiality to all participants, promising not to disclose any identifying information that was gathered (Creswell, 2007).

**Conclusion**

As mentioned previously the most important factor in student achievement is the quality of the teacher (Marzano, 2003; Hanushek 2009; Rothstein, 2004). Furthermore, Performance Improvement Theory tells us that improvement of individual performance is dependent upon clear goals and effective feedback on performance in order for individuals to change their behaviors and performance (Swanson, 1999; Smither, London, and Reilly, 2005). This study explored whether teachers who had experienced the new Massachusetts evaluation system viewed it as changing their instructional practice. Positive improvements in teaching practice will increase teacher quality and student achievement. If educators do not see the new evaluation system as improving practice, then the system should be adjusted to ensure this essential aspect is being achieved.

**Chapter IV: Research Findings**

**Reporting of the Findings and Analysis**

Chapter four presents the findings and analysis of this qualitative study starting with a description of the study’s context, including a description of the participants, followed by a presentation of the findings. The findings will be presented in response to each research questions, which are as follows:

1. Do teachers perceive the new Massachusetts teacher evaluation system as a valid and useful tool for changing their teaching practice?
2. How do teachers perceive the evaluation process of self-reflection, goal setting, and evaluation using the rubric as a valid and worthwhile assessment of effective teaching?

3. Do teachers perceive their use of the tool and engagement in the process as an effective means toward changing their teaching practice for greater student achievement?

Survey responses and an analysis of focus groups and interviews will be presented at it pertains to each research question.

**Study Context**

The purpose of this study was to examine the perceptions of teachers during the pilot of the Massachusetts evaluation model as required by the “Race to the Top” grant in three of the eleven early adopter districts. These districts volunteered to implement the new developing evaluation model, provided information to the Department of Elementary and Secondary Education to assist in the development and future implementation of the new evaluation system. The examination of these results will provide information on the perceived validity of the instruments and evaluation process and if, in the teachers view, the implementation of the instruments had the potential to cause changes and improvements in teaching practices.

**Survey Participants**

Table 1 provides the number and percentage of participants by number of years teaching, grade level teaching at the time of the survey, and the district in which they teach.
There were 52 participants in the survey portion of this study. All participants were certified school personnel, including two guidance counselors who participated in their district’s pilot of the new Massachusetts teacher evaluation system. Participants from District A composed the majority of the middle school respondents. Conversely, those in District B constituted the bulk of the responses from high school educators. Table 1 presents the number of participants by district, grade level of school, and number of years teaching. Due to the low numbers of participants at the elementary level (n=2), this study will not attempt to draw conclusions of the perceptions of teachers who piloted the new evaluation system in elementary schools. Furthermore, due to the low number of participants in the pilot in district C (n=4) and the fact that only one person per school piloted the new evaluation system, this study will not
attempt to draw separate conclusions from this group, but their responses were utilized in the overall analysis

**Survey Analysis**

The survey data was first compiled and disaggregated by length of time teaching, grade level of the school they currently teach in, and their district. The data was then organized into distribution tables where the mean and standard deviation (σ) was calculated and the percentages of responses in each of the Likert categories were calculated. In order to determine if the answers represented a normal distribution at the p< 0.05 level, the Fisher Exact test was conducted for each question and subgroup.

**Focus Groups and Interview Participants**

Respondents were invited to participate in focus groups held in each participating district. In total, nine educators agreed to participate in the focus groups, which in some cases became simply an individual interview when only one volunteered. Participants included one elementary teacher in an urban district; five middle school teachers from two different suburban districts, and three were high school educators from one suburban district. In each focus group or individual interview, the process and questions were consistent; the responses were recorded and then transcribed for analysis. To ensure accuracy of the interviews a second meeting was held with the participants to verify the validity of the transcriptions.

**Focus Group and Interview Data**

In order to analyze the qualitative data from the focus groups and interviews, participant responses were first transcribed then analyzed for themes and clusters of answer types. The responses were then categorized and used to add insight and commentary to survey question results. As stated previously, the researcher examined this data to identify significant statements,
clusters of meaning, and subsequently themes reflecting the teachers’ feelings and perceptions of the new evaluation system’s impact on their teaching practice and self-reflection, and its ability to evaluate effective teaching.

**Research Question #1: Do Teachers Perceive the New Massachusetts Teacher Evaluation System as a Valid and Useful Tool for Changing their Teaching Practice?**

To address this research question this section will review participants’ responses to questions about SMART goals, the overall effect of the new teacher evaluation process on teachers’ professional practice, if the evaluator was able to collect sufficient evidence in the evaluation process, and if participating in the new evaluation process positively impacted their classroom teaching methods.

**Did SMART goals, as part of the new teacher evaluation process, help teachers focus on the improvement of their teaching practice?** The new teacher evaluation process requires that teachers develop goals that are strategic, measurable, action orientated, rigorous but realistic, and timed and tracked in the areas of student achievement and improving teacher practice. Following a process of self-reflection on student achievement and professional practice as related to the teacher evaluation rubric, teachers were then asked to develop SMART Goals. To assess the degree to which teachers using the evaluation system agreed with the following statement the following question was posed; Setting SMART goals has helped me focus on improving my teaching practice.
Table 2

*SMART goals helped focus teachers on improving their teaching practice.*

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither</th>
<th>Agree</th>
<th>Strongly Agree</th>
<th>Rating Average</th>
<th>St. Dev.</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5 Years</td>
<td>0%</td>
<td>20%</td>
<td>20%</td>
<td>60%</td>
<td>0%</td>
<td>3.40</td>
<td>0.89</td>
<td>5</td>
</tr>
<tr>
<td>6-10 Years</td>
<td>6%</td>
<td>29%</td>
<td>18%</td>
<td>47%</td>
<td>0%</td>
<td>3.06</td>
<td>1.03</td>
<td>17</td>
</tr>
<tr>
<td>11-15 Years</td>
<td>0%</td>
<td>22%</td>
<td>11%</td>
<td>67%</td>
<td>0%</td>
<td>3.44</td>
<td>0.88</td>
<td>9</td>
</tr>
<tr>
<td>&gt;15 Years</td>
<td>0%</td>
<td>14%</td>
<td>29%</td>
<td>43%</td>
<td>14%</td>
<td>3.57</td>
<td>0.88</td>
<td>21</td>
</tr>
<tr>
<td>Elementary School</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>50%</td>
<td>50%</td>
<td>4.50</td>
<td>0.71</td>
<td>2</td>
</tr>
<tr>
<td>Middle School</td>
<td>0%</td>
<td>5%</td>
<td>19%</td>
<td>67%</td>
<td>10%</td>
<td>3.81^</td>
<td>0.68</td>
<td>21</td>
</tr>
<tr>
<td>High School</td>
<td>3%</td>
<td>34%</td>
<td>24%</td>
<td>38%</td>
<td>0%</td>
<td>2.97^</td>
<td>0.94</td>
<td>29</td>
</tr>
<tr>
<td>District A</td>
<td>0%</td>
<td>8%</td>
<td>20%</td>
<td>64%</td>
<td>8%</td>
<td>3.72+</td>
<td>0.74</td>
<td>25</td>
</tr>
<tr>
<td>District B</td>
<td>4%</td>
<td>39%</td>
<td>26%</td>
<td>30%</td>
<td>0%</td>
<td>2.83+</td>
<td>0.94</td>
<td>23</td>
</tr>
<tr>
<td>District C</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>75%</td>
<td>25%</td>
<td>4.25</td>
<td>0.50</td>
<td>4</td>
</tr>
<tr>
<td>Dist. A-M.S.</td>
<td>0%</td>
<td>5%</td>
<td>20%</td>
<td>65%</td>
<td>10%</td>
<td>3.80#</td>
<td>0.70</td>
<td>20</td>
</tr>
<tr>
<td>Dist. B-H.S.</td>
<td>4%</td>
<td>39%</td>
<td>26%</td>
<td>30%</td>
<td>0%</td>
<td>2.83#</td>
<td>0.94</td>
<td>23</td>
</tr>
<tr>
<td>Non-Dist. B-H.S.</td>
<td>0%</td>
<td>17%</td>
<td>17%</td>
<td>67%</td>
<td>0%</td>
<td>3.50</td>
<td>0.84</td>
<td>6</td>
</tr>
<tr>
<td>OVERALL</td>
<td>2%</td>
<td>21%</td>
<td>21%</td>
<td>50%</td>
<td>6%</td>
<td>3.37</td>
<td>0.95</td>
<td>52</td>
</tr>
</tbody>
</table>

^ Middle school teachers agreed significantly more positive than high school teachers (Fisher Exact Test, p=0.010)

^ Teachers from District A were overall in much more agreement with the survey statement than teachers from District B (Fisher Exact Test p=0.004)

# Middle school teachers from District A responded more positively than high school teachers from District B (Fisher Exact Test p=0.004)

As can be seen in Table 2, the overall mean teacher response rating for this question was 3.37 (σ = 0.95), with just slightly more than the majority of the respondents, 56% responding that they agree or strongly agree and 23% disagreeing or strongly disagreeing. In fact, the majority
of every subgroup agreed or strongly agreed to this question except for the teachers in District B high school, and those who had been teaching for six to ten years (although close to half of these teachers, 47%, agreed). District A and the District A middle school teachers in particular provided a fairly positive response to this question with an average of 3.80 and 3.81 respectively, with 75% and 77% agreeing or strongly agreeing, respectively. The 6-10 years of teaching experience participants were in less agreement than other years of teaching experience groups (47%). The middle school teachers were in greater agreement than the high school teachers. Furthermore, teacher participants in District A demonstrated more agreement than District B teacher participants. In using the Fisher Exact test to determine significant differences between groups, the Fisher Exact disclosed a significant difference between the middle and high school teacher participants (p=0.01), and District A and District B teachers (p=0.004).

Many teacher responses in the focus groups complemented the finding that the SMART goal process positively influenced their teaching practice. One teacher put it well when she shared her experience as follows:

Had I not had that self-assessment piece I probably would have stayed in that ho hum routine survival mode, get through. But because I had to stop and think about why am I teaching what am I trying to get across to my students, made me stop and think and then better the lesson, (to get) it across to them better.

In a different interview session another teacher was quite reflective on the process, stating “I was reflecting more like (she) said and I was stepping back all the time kept reflecting and kept reflecting, obviously you (will) have a better lesson.” Of course, not all of teachers’ responses were in agreement with this sentiment, given the results of the survey. One teacher shared,
I wish I had received a little bit more guidance because I see part of the problem now was I had a lot of goals, but they have to change, because they were too much out of my control, at least the student-learning goal, too much out of my control. They’re too hard.

I had terrible measurements for them.

This statement illustrates the learning curve involved for some teachers in the new goal setting process.

The new evaluation process resulted in changes in my teaching methods. This question had to do with whether the new evaluation process resulted in changes in their teaching methods. The new evaluation system requires teachers to self-reflect, using the teacher evaluation rubric, on their teacher practice to find areas were growth is needed, and then asks teachers to develop SMART goals to foster growth in these areas. This question asks participants if this process has resulted in any change in teaching methods. Furthermore, this question is central in disclosing whether teachers perceive the new Massachusetts teacher evaluation system as a valid and useful tool for changing their teaching practice. The responses by participants across the various categories of this analysis are presented in Table 3.
**Table 3**

*The new evaluation process changed my teaching methods*

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither</th>
<th>Agree</th>
<th>Strongly Agree</th>
<th>Rating Average</th>
<th>St. Dev.</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5 Years</td>
<td>0%</td>
<td>40%</td>
<td>20%</td>
<td>40%</td>
<td>0%</td>
<td>3.00</td>
<td>1.00</td>
<td>5</td>
</tr>
<tr>
<td>6-10 Years</td>
<td>6%</td>
<td>41%</td>
<td>18%</td>
<td>35%</td>
<td>0%</td>
<td>2.82</td>
<td>1.01</td>
<td>17</td>
</tr>
<tr>
<td>11-15 Years</td>
<td>11%</td>
<td>78%</td>
<td>0%</td>
<td>11%</td>
<td>0%</td>
<td>2.11*</td>
<td>0.78</td>
<td>9</td>
</tr>
<tr>
<td>&gt;15 Years</td>
<td>5%</td>
<td>14%</td>
<td>29%</td>
<td>48%</td>
<td>5%</td>
<td>3.33*</td>
<td>0.97</td>
<td>21</td>
</tr>
</tbody>
</table>

|               |                   |          |         |       |               |              |          |                |
| Elementary School | 50%             | 0%      | 0%      | 50%   | 0%            | 2.50         | 2.12     | 2              |
| Middle School   | 0%                | 19%      | 24%     | 52%   | 5%            | 3.43^        | 0.87     | 21             |
| High School     | 7%                | 52%      | 17%     | 24%   | 0%            | 2.59^        | 0.95     | 29             |

|               |                   |          |         |       |               |              |          |                |
| District A     | 0%                | 20%      | 25%     | 50%   | 5%            | 3.40+        | 0.86     | 25             |
| District B     | 9%                | 57%      | 13%     | 22%   | 0%            | 2.48+        | 0.95     | 23             |
| District C     | 25%               | 25%      | 0%      | 50%   | 0%            | 2.75         | 1.50     | 4              |

|               |                   |          |         |       |               |              |          |                |
| Dist. A M.S.  | 0%                | 20%      | 25%     | 50%   | 5%            | 3.40#        | 0.88     | 20             |
| Dist. B H.S.  | 9%                | 57%      | 13%     | 22%   | 0%            | 2.48#        | 0.95     | 23             |
| Non-dist. B-  | 0%                | 33%      | 33%     | 33%   | 0%            | 3.00         | 0.89     | 6              |
| H. S.         |                   |          |         |       |               |              |          |                |
| OVERALL       | 6%                | 37%      | 19%     | 37%   | 2%            | 2.92         | 1.02     | 52             |

* Teachers with 11-15 year of teaching answered significantly less positive than teachers with over 15 years of teaching experience (Fisher Exact Test, p=0.049)

^ Middle school teachers agreed significantly more positive than high school teachers (Fisher Exact Test, p=0.022)

* Teachers from District A were overall in much more agreement with the survey statement than teachers from District B (Fisher Exact Test p=0.032)

# Middle school teachers from District A responded more positively than high school teachers from District B (Fisher Exact Test p=0.026)

As can be seen in Table 3, overall the teachers participating in this survey averaged slightly less than 3.0 (2.92, σ = 1.02) with only 37% agreeing and 2% strongly agreeing, and 37% disagreeing and 6% strongly disagreeing. The only groups that averaged above 3.0 in response to this question were the teachers with more than 15 years of teaching experience,
middle school teachers, middle school teachers in District A, and all teachers in District A. In these four groups 53%, 57%, 55%, and 55% respectively agreed or strongly agreed. In contrast, less than 40% of all teachers in the other groups agreed or strongly agreed. The 11-15 years of teaching experience group averaged only 2.11 and had only 11% agree and no one strongly agreeing with the statement.

When testing for significant differences between the groups, teachers with 11-15 years of experience responded significantly less in agreement (2.11) than those with more than 15 years of teaching experience (3.33) (Fisher Exact Test, p=0.049). It is also notable that middle school teachers from District A responded more positively than high school teachers in District B (Fisher Exact Test, p=0.004). In addition, and not different from the previous question, District A teachers were far more likely to agree than District B teachers (Fisher Exact Test, p=0.031).

Of those teachers participating in the focus groups, the focus group participants expressed that the evaluation process did influence their teaching practice. One teacher stated, “I would have to say it’s because of that (evaluation process), I am better than I was”. Another noted I think I was a little more aware that at any time someone could pop in, so I wanted to make sure that what I was doing was, not perfect, but what I said I was going to be doing. I think it made me a little more organized. It made me look for better lessons for my kids, more game-like, more things that they were going to really enjoy. Yeah, it definitely, I think it definitely made me better, and a lot more aware than I was in the past.

It must be noted that although others mentioned changes in their practice, the examples were referencing minutia, and not overarching teaching practice. For example, one teacher stated, “He [the evaluating administrator] made some critiques of the PowerPoint’s that I used, basically,
you know too much text on one slide, and I decided he was correct, I re-did the Power Points.” Although the teacher made changes in a presentation method, it was unclear if the entire message for improvement was conveyed or received.

My evaluator was able to gather sufficient evidence to accurately assess my performance. The next question asked if teachers felt that their evaluator was able to gather sufficient evidence to assess accurately their performance. In the Massachusetts model evaluation system a dramatic shift was made from pre-scheduled observations to unannounced and often shorter walk through observations, the use of the extensive teacher evaluation rubric, and artifacts of evidence presented by the teachers (Massachusetts Department of Elementary and Secondary Education, 2012). These expectations represent an increased in the evidence and assessment of teacher performance from previous evaluation systems. Responses by teachers across the various categories of this analysis are presented in Table 4.
Table 4

Evaluators were able to gather sufficient evidence to assess accurately teacher performance.

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither</th>
<th>Agree</th>
<th>Strongly Agree</th>
<th>Rating Average</th>
<th>St. Dev.</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5 Years</td>
<td>0%</td>
<td>20%</td>
<td>0%</td>
<td>60%</td>
<td>20%</td>
<td>3.80</td>
<td>0.02</td>
<td>5</td>
</tr>
<tr>
<td>6-10 Years</td>
<td>0%</td>
<td>12%</td>
<td>35%</td>
<td>53%</td>
<td>0%</td>
<td>3.41*</td>
<td>0.71</td>
<td>17</td>
</tr>
<tr>
<td>11-15 Years</td>
<td>0%</td>
<td>0%</td>
<td>22%</td>
<td>78%</td>
<td>0%</td>
<td>3.78*</td>
<td>0.44</td>
<td>9</td>
</tr>
<tr>
<td>&gt;15 Years</td>
<td>5%</td>
<td>5%</td>
<td>0%</td>
<td>71%</td>
<td>19%</td>
<td>3.95</td>
<td>0.92</td>
<td>21</td>
</tr>
<tr>
<td>Elementary School</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>50%</td>
<td>50%</td>
<td>4.50</td>
<td>0.71</td>
<td>2</td>
</tr>
<tr>
<td>Middle School</td>
<td>0%</td>
<td>10%</td>
<td>5%</td>
<td>76%</td>
<td>10%</td>
<td>3.86</td>
<td>0.73</td>
<td>21</td>
</tr>
<tr>
<td>High School</td>
<td>4%</td>
<td>9%</td>
<td>30%</td>
<td>48%</td>
<td>9%</td>
<td>3.59</td>
<td>0.87</td>
<td>29</td>
</tr>
<tr>
<td>District A</td>
<td>0%</td>
<td>8%</td>
<td>4%</td>
<td>80%</td>
<td>8%</td>
<td>3.88^</td>
<td>0.67</td>
<td>25</td>
</tr>
<tr>
<td>District B</td>
<td>4%</td>
<td>9%</td>
<td>30%</td>
<td>48%</td>
<td>9%</td>
<td>3.48^</td>
<td>0.95</td>
<td>23</td>
</tr>
<tr>
<td>District C</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>75%</td>
<td>25%</td>
<td>4.25</td>
<td>0.50</td>
<td>4</td>
</tr>
<tr>
<td>Dist. A -M.S.</td>
<td>0%</td>
<td>10%</td>
<td>5%</td>
<td>75%</td>
<td>10%</td>
<td>3.85#</td>
<td>0.77</td>
<td>20</td>
</tr>
<tr>
<td>Dist. B- H.S.</td>
<td>4%</td>
<td>9%</td>
<td>30%</td>
<td>48%</td>
<td>9%</td>
<td>3.48++</td>
<td>0.95</td>
<td>23</td>
</tr>
<tr>
<td>Non Dist. B- H.S.</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>100%</td>
<td>0%</td>
<td>4.00+</td>
<td>0.00</td>
<td>6</td>
</tr>
<tr>
<td>OVERALL</td>
<td>2%</td>
<td>8%</td>
<td>15%</td>
<td>65%</td>
<td>10%</td>
<td>3.73</td>
<td>0.82</td>
<td>52</td>
</tr>
</tbody>
</table>

* Teachers with 6-10 year of teaching answered significantly less positive than teachers with over 15 years of teaching experience (Fisher Exact Test, p=0.012)

^ Teachers from District A were overall in much more agreement with the survey statement than teachers from District B (Fisher Exact Test p=0.012)

+ High School Teachers who are not in District B were significantly more in agreement that high school teachers in District B (Fisher Exact Test p=0.049)

#District A Middle School teachers significantly responded more positively than District B high school teachers (Fisher Exact Test p=0.043)

As can be seen in Table 4, the overall response average was relatively high, 3.73 (σ = 0.82), with 65% agreeing and 10% strongly agreeing with the statement that the evaluator was able to collect sufficient evidence to assess teacher performance. In fact, every subgroup averaged above 3.4 if not well above 3.4 and over 50% if not much greater either agreed or
strongly agreed with the statement. The lowest reporting group was the teachers with six to ten years of teaching experience, with an average response of 3.41 average with 53% agreeing and 0% strongly agreeing. In contrast, the District A teachers’ response averaged 3.88 with 80% agreeing and 8% strongly agreeing. In comparing the groups to determine if there was any significant differences, teachers with over 15 years of teaching experience agreed more so than those with 6-10 years of experience teaching (Fisher Exact Test, p=0.012). And, in keeping with previous questions, District A teachers agreed more than teachers from District B, but the difference was not as great (3.88 vs. 3.48, respectively; Fisher Exact Test, p=0.012). Overall, and in contrast to the other questions on the survey, more teachers across all subgroups agreed or strongly agreed than disagreed or strongly disagreed.

Regarding this question, the focus group participants shared a variety of opinions of the process. One shared, “I think with the new system even if they come through when you are watching a movie, doesn’t matter because they are just out and about so much more that they know what is going on in the classroom.” Another shared a less positive perspective, stating, “I think that it’s difficult to assess my teaching performance, and what happens in my classroom throughout the year with a couple of pieces of evidence to support each indicator and those informal observations. It’s (only) a snapshot of what I do.” Most respondents however expressed more positive sentiments towards the process. For example, “I love the walk throughs as opposed to putting on a dog and pony show for one day.” Conversely, a few were skeptical of the ability to collect enough evidence, “I think that it’s difficult to assess my teaching performance, and what happens in my classroom throughout the year with a couple of pieces of evidence to support each indicator and with, you know, those informal observations.”
Looking back, the evaluation process resulted in positive changes in my professional practice. The next question asked teachers to look back at the new evaluation process and share if they thought it resulted in positive changes in their professional practice. As one will recall, the goal of the new evaluation process is to provide a focus for the improvement of teaching practices and ensure that only highly effective individuals lead instruction in our classrooms (Rothstein, 2010; Massachusetts Department of Elementary and Secondary Education, 2011). The responses by participants across the various categories of this analysis are presented in Table 5.
Table 5

The evaluation process results in positive changes in professional practice.

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither</th>
<th>Agree</th>
<th>Strongly Agree</th>
<th>Rating Average</th>
<th>St. Dev.</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5 Years</td>
<td>0%</td>
<td>20%</td>
<td>20%</td>
<td>60%</td>
<td>0%</td>
<td>3.40</td>
<td>0.89</td>
<td>5</td>
</tr>
<tr>
<td>6-10 Years</td>
<td>6%</td>
<td>35%</td>
<td>18%</td>
<td>41%</td>
<td>0%</td>
<td>2.94</td>
<td>1.03</td>
<td>17</td>
</tr>
<tr>
<td>11-15 Years</td>
<td>11%</td>
<td>11%</td>
<td>11%</td>
<td>67%</td>
<td>0%</td>
<td>3.33</td>
<td>1.08</td>
<td>9</td>
</tr>
<tr>
<td>&gt;15 Years</td>
<td>5%</td>
<td>14%</td>
<td>19%</td>
<td>48%</td>
<td>14%</td>
<td>3.52</td>
<td>1.08</td>
<td>21</td>
</tr>
<tr>
<td>Elementary School</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>50%</td>
<td>50%</td>
<td>4.50</td>
<td>0.71</td>
<td>2</td>
</tr>
<tr>
<td>Middle School</td>
<td>0%</td>
<td>5%</td>
<td>14%</td>
<td>76%</td>
<td>5%</td>
<td>3.81*</td>
<td>0.60</td>
<td>21</td>
</tr>
<tr>
<td>High School</td>
<td>10%</td>
<td>34%</td>
<td>21%</td>
<td>31%</td>
<td>3%</td>
<td>2.83*</td>
<td>1.10</td>
<td>29</td>
</tr>
<tr>
<td>District A</td>
<td>0%</td>
<td>16%</td>
<td>16%</td>
<td>64%</td>
<td>4%</td>
<td>3.56*</td>
<td>0.82</td>
<td>25</td>
</tr>
<tr>
<td>District B</td>
<td>13%</td>
<td>30%</td>
<td>22%</td>
<td>30%</td>
<td>4%</td>
<td>2.83*</td>
<td>1.15</td>
<td>23</td>
</tr>
<tr>
<td>District C</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>75%</td>
<td>25%</td>
<td>4.25</td>
<td>0.50</td>
<td>4</td>
</tr>
<tr>
<td>Dist. A-M.S.</td>
<td>0%</td>
<td>5%</td>
<td>15%</td>
<td>75%</td>
<td>5%</td>
<td>3.80^</td>
<td>0.62</td>
<td>20</td>
</tr>
<tr>
<td>Dist. B-H.S.</td>
<td>13%</td>
<td>30%</td>
<td>22%</td>
<td>30%</td>
<td>4%</td>
<td>2.83^</td>
<td>1.15</td>
<td>23</td>
</tr>
<tr>
<td>Non Dist.B-H.S.</td>
<td>0%</td>
<td>50%</td>
<td>17%</td>
<td>33%</td>
<td>0%</td>
<td>2.83</td>
<td>0.98</td>
<td>6</td>
</tr>
<tr>
<td>OVERALL</td>
<td>6%</td>
<td>21%</td>
<td>17%</td>
<td>50%</td>
<td>6%</td>
<td>3.29</td>
<td>1.05</td>
<td>52</td>
</tr>
</tbody>
</table>

+ Middle school teachers were significantly more positive than high school teachers (p=0.001)
* Teachers from District A were overall in much more agreement with the survey statement than teachers from District B (Fisher Exact Test p=0.014)
^ District A middle school teachers significantly responded more positively than District B high school teachers (Fisher Exact Test p=0.002)

As can be seen in Table 5, more teachers agreed with the statement that the evaluation system resulted in positive changes in their professional practice (56%) than those who disagreed (27%), with 17% neither agreeing or disagreeing. Demonstrating an ongoing trend with the 6-10 years of teaching experience group, the high school teachers, and teachers in District B, more
teachers disagreed than agreed, with the average mean response falling below 3. Only 34% of High School and District B teachers agreed or strongly agree.

Although there was no statistically significant difference between the years of teaching experience groups on this question, the overall trend continued in agreement or lack of agreement across other groups. Once again District A middle school teachers responded significantly higher than the District B high school teachers, with 80% of the district A middle school teachers agreeing or strongly agreeing vs. only 34% of district B high school teachers (Fisher Exact Test, p=0.002). Overall, again District A teachers are significantly more positive than District B teachers (Fisher Exact Test, p=0.014), as were all of the middle school teachers in contrast to all of the high school teachers (p=0.001).

In the focus groups, the majority of teachers reported that the new system did cause positive changes in their teaching practice. As one of the teachers stated,

One of our goals for the high school was on literacy so we had a whole new rubric for writing for grading student writing. I incorporated this into my labs. I know I am reading the lab conclusions differently than before, and I think it is a better way, rather than just nuts and bolts, reading whole argument, much more logically, does it have a more logical flow to it, that is definitely one thing that I have improved based on (the evaluation system).

Another teacher succinctly said, “I would have to say because of that [the evaluation system] I am better than I was.” Others had a more measured or mixed reaction to questions regarding changes in their practice. As one teacher said, “There were a couple of things I took from it that were great, and then a couple of things where you were speaking different languages entirely.” In general, the responses proved to be mixed across the groups.
Summary of results to research question one. Overall, when looking at whether teachers viewed the new evaluation system in Massachusetts as a valid and useful tool in changing their teaching practice, the overall results were quite mixed. Teachers’ moderately agreed that SMART goals helped them to improve their teaching practice ($\mu=3.37$, $\sigma=.95$), as well as moderately agreed that the process resulted in positive changes to their professional practice ($\mu=3.29$, $\sigma=1.05$). More positive was their response to the question of whether their evaluators were able to collect enough evidence to accurately assess their performance ($\mu=3.73$, $\sigma=.82$). However, responses to the question of whether the evaluation process changed their teaching practice the responses were much more mixed, with only 39% agreeing or strongly agreeing and 43% disagreeing or strongly disagreeing ($\mu=2.92$, $\sigma=1.02$).

These responses are somewhat underwhelming given the overarching purpose of the new teacher evaluation system to improve teacher quality and teachers’ practice. In the focus group discussions participants shared that they did experience some changes in their practice but the examples were relatively minor in scope. Important to note, however, is the fact that two diametrically opposed perceptions emerged between the District A middle school and District C participants and the District A middle school and District C participants. The District A middle school and District C participants consistently reported more positive perceptions regarding the teacher evaluation system as an impetus and tool for reflecting upon and improving their practice than District B participants, who happened to be all high school teachers.
Research Question #2: How do Teachers Perceive the Evaluation Process of Self-reflection, Goal setting, and Evaluation using the Rubric as Valid and Worthwhile Assessment of Effective Teaching?

The second research question examines teachers’ perceptions of the Massachusetts Evaluation system as it relates to effective teaching.

The self-reflection process using the rubric caused serious consideration of individual’s teaching practice. The self-reflection process requires teachers to reflect on their teaching methods and practice in reference to the standards and indicators of effective teaching as described in the teacher evaluation rubric. This effort is then used in the development of SMART goals for the purpose of improving one’s practice and fostering effective teaching (Massachusetts Department of Elementary and Secondary Education, 2011). Teachers’ responses to the question of whether the self-reflection process cause serious consideration of their teaching practice is presented in Table 6.
Table 6

The self-reflection process caused serious consideration of individual’s teaching practice

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither</th>
<th>Agree</th>
<th>Strongly Agree</th>
<th>Rating Average</th>
<th>St. Dev.</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5 Years</td>
<td>0%</td>
<td>40%</td>
<td>20%</td>
<td>40%</td>
<td>0%</td>
<td>3.00</td>
<td>1.00</td>
<td>5</td>
</tr>
<tr>
<td>6-10 Years</td>
<td>6%</td>
<td>18%</td>
<td>18%</td>
<td>47%</td>
<td>12%</td>
<td>3.41</td>
<td>1.01</td>
<td>17</td>
</tr>
<tr>
<td>11-15 Years</td>
<td>0%</td>
<td>11%</td>
<td>0%</td>
<td>89%</td>
<td>0%</td>
<td>3.78</td>
<td>0.67</td>
<td>9</td>
</tr>
<tr>
<td>&gt;15 Years</td>
<td>0%</td>
<td>14%</td>
<td>24%</td>
<td>43%</td>
<td>19%</td>
<td>3.67</td>
<td>0.97</td>
<td>21</td>
</tr>
<tr>
<td>Elementary School</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>50%</td>
<td>50%</td>
<td>4.50</td>
<td>0.71</td>
<td>2</td>
</tr>
<tr>
<td>Middle School</td>
<td>0%</td>
<td>5%</td>
<td>10%</td>
<td>71%</td>
<td>14%</td>
<td>3.95+</td>
<td>0.67</td>
<td>21</td>
</tr>
<tr>
<td>High School</td>
<td>3%</td>
<td>28%</td>
<td>24%</td>
<td>38%</td>
<td>7%</td>
<td>3.17+</td>
<td>1.03</td>
<td>29</td>
</tr>
<tr>
<td>District A</td>
<td>0%</td>
<td>5%</td>
<td>10%</td>
<td>70%</td>
<td>15%</td>
<td>3.95^</td>
<td>0.83</td>
<td>25</td>
</tr>
<tr>
<td>District B</td>
<td>4%</td>
<td>26%</td>
<td>26%</td>
<td>35%</td>
<td>9%</td>
<td>3.17^</td>
<td>1.03</td>
<td>23</td>
</tr>
<tr>
<td>District C</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>75%</td>
<td>25%</td>
<td>4.25</td>
<td>0.50</td>
<td>4</td>
</tr>
<tr>
<td>District A- M.S.</td>
<td>0%</td>
<td>5%</td>
<td>10%</td>
<td>70%</td>
<td>15%</td>
<td>3.95#</td>
<td>0.64</td>
<td>20</td>
</tr>
<tr>
<td>District B - H.S.</td>
<td>4%</td>
<td>26%</td>
<td>26%</td>
<td>35%</td>
<td>9%</td>
<td>3.17#</td>
<td>1.07</td>
<td>23</td>
</tr>
<tr>
<td>Non-Dist B-H.S.</td>
<td>0%</td>
<td>8%</td>
<td>17%</td>
<td>50%</td>
<td>0%</td>
<td>3.17</td>
<td>0.98</td>
<td>6</td>
</tr>
<tr>
<td>OVERALL</td>
<td>2%</td>
<td>17%</td>
<td>17%</td>
<td>52%</td>
<td>12%</td>
<td>3.54</td>
<td>0.98</td>
<td>52</td>
</tr>
</tbody>
</table>

^ Teachers from District A were overall in much more agreement with the survey statement than teachers from District B (Fisher Exact Test p=0.021)
+ Middle school Teachers were significantly more in agreement than high school teachers in (Fisher Exact Test p=0.003)
#District A Middle School teachers significantly responded more positively than District B high school teachers (Fisher Exact Test p=0.043)

As can be seen in Table 6, the majority of teachers (64%) perceived the self-evaluation process as having influenced their reflection on their teaching practice, with an average response of 3.54. By category, the groups responding most positively to this item were the middle school teachers in District A, with 70% agreeing and 15% strongly agreeing (85% overall). Lower
responses were found in the one to five years of experience teachers, with only two of the five teachers (40%) in this category agreeing, high school teachers, with only 45% agreeing, and district B teachers, with only 41% agreeing. Once again there were significant differences found between the middle school and high school groups (Fisher Exact Test, p=0.003), District A and District B teachers (Fisher Exact Test, p=0.021), and District A middle school and District B high school teachers (Fisher Exact Test, p=0.005).

In the focus groups, this researcher found that through their preparation programs, early career teachers had been indoctrinated in the self-reflection process, and thus this form of reflection was not a new experience. One third-year teacher for example shared, “I had to do that for my undergrad and graduate (degrees).” In addition, District A teachers had previously been using a reflection and goal setting system as part of their past, district-based evaluation, which may have contributed to their fairly positive response to the new evaluation system in contrast to others. As one participant put it, “We were already ahead of the game, we had already had walk-throughs, already did goal setting and Smart Goals.” It is important to note that in District B the new evaluation system process was very different to their previous evaluation processes, which may have contributed to their lower responses given it was a new activity. The newness and uncertainty was illustrated when one teacher shared, “I don’t know, you know, maybe in another year I’ll be singing its praises.”

The rubric, the standards and indicators of effective teaching practice, is a comprehensive guide of the traits of effective teaching. This survey question was designed to examine teacher perceptions of the rubric as a representation of the standards of effective teaching. The teacher evaluation rubric was designed to define the key the components of effective teaching and through defining them to facilitate improved teaching practice in the
Commonwealth of Massachusetts (Danielson, 2007; Massachusetts Department of Elementary and Secondary Education, 2011). Responses by teachers across the various subgroups in response to the question of whether they consider the rubric to be a comprehensive guide to the traits of effective teaching are presented in Table 7.

Table 7

*The rubric is a comprehensive representation of effective teaching practice.*

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither</th>
<th>Agree</th>
<th>Strongly agree</th>
<th>Rating Average</th>
<th>St. Dev.</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5 Years</td>
<td>0%</td>
<td>20%</td>
<td>20%</td>
<td>60%</td>
<td>0%</td>
<td>3.40</td>
<td>0.89</td>
<td>5</td>
</tr>
<tr>
<td>6-10 Years</td>
<td>0%</td>
<td>12%</td>
<td>6%</td>
<td>82%</td>
<td>0%</td>
<td>3.71</td>
<td>0.69</td>
<td>17</td>
</tr>
<tr>
<td>11-15 Years</td>
<td>5%</td>
<td>10%</td>
<td>14%</td>
<td>57%</td>
<td>14%</td>
<td>3.67</td>
<td>0.92</td>
<td>9</td>
</tr>
<tr>
<td>&gt;15 Years</td>
<td>0%</td>
<td>14%</td>
<td>24%</td>
<td>43%</td>
<td>19%</td>
<td>3.67</td>
<td>0.97</td>
<td>21</td>
</tr>
<tr>
<td>Elementary School</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>50%</td>
<td>50%</td>
<td>4.50</td>
<td>0.71</td>
<td>2</td>
</tr>
<tr>
<td>Middle School</td>
<td>0%</td>
<td>0%</td>
<td>10%</td>
<td>86%</td>
<td>5%</td>
<td>3.95^</td>
<td>0.38</td>
<td>21</td>
</tr>
<tr>
<td>High School</td>
<td>5%</td>
<td>30%</td>
<td>15%</td>
<td>59%</td>
<td>10%</td>
<td>3.45^</td>
<td>1.02</td>
<td>29</td>
</tr>
<tr>
<td>District A</td>
<td>0%</td>
<td>0%</td>
<td>5%</td>
<td>90%</td>
<td>5%</td>
<td>4.00</td>
<td>0.74</td>
<td>25</td>
</tr>
<tr>
<td>District B</td>
<td>4%</td>
<td>13%</td>
<td>9%</td>
<td>65%</td>
<td>9%</td>
<td>3.61</td>
<td>0.99</td>
<td>23</td>
</tr>
<tr>
<td>District C</td>
<td>0%</td>
<td>0%</td>
<td>25%</td>
<td>50%</td>
<td>25%</td>
<td>4.00</td>
<td>0.82</td>
<td>4</td>
</tr>
<tr>
<td>Dist. A-M.S.</td>
<td>0%</td>
<td>0%</td>
<td>5%</td>
<td>90%</td>
<td>5%</td>
<td>4.00</td>
<td>0.32</td>
<td>20</td>
</tr>
<tr>
<td>Dist. B-H.S.</td>
<td>4%</td>
<td>13%</td>
<td>9%</td>
<td>65%</td>
<td>9%</td>
<td>3.61</td>
<td>0.99</td>
<td>23</td>
</tr>
<tr>
<td>Non-Dist.B-H.S.</td>
<td>0%</td>
<td>50%</td>
<td>17%</td>
<td>33%</td>
<td>0%</td>
<td>2.83</td>
<td>0.98</td>
<td>6</td>
</tr>
<tr>
<td>OVERALL</td>
<td>3%</td>
<td>21%</td>
<td>17%</td>
<td>69%</td>
<td>14%</td>
<td>3.69</td>
<td>0.85</td>
<td>52</td>
</tr>
</tbody>
</table>

^ Middle school Teachers were significantly more in agreement than high school teachers in (Fisher Exact Test p=0.03)
As can be seen in Table 7 teachers overall in these pilot districts perceived the teacher evaluation rubric to be for the most part representative of effective teaching practice (3.69, σ =0.85). All subgroups had an average response greater than 3.0 with many subgroups responding much greater than that. The lower response groups included high school teachers and those in their early careers, one to five years. When testing for significance, and breaking with earlier trends, there was only one set of groups that exhibited a significant difference, and that was the middle school teachers from the high school teachers, with 91% of middle school teachers agreeing versus 69% of high school teachers agreeing (Fisher Exact Test, p=0.030). Even so, still 69% of agreement is relatively strong.

In the follow up focus groups, many stated that the rubric was very comprehensive, with one teacher commenting, “They’re good standards in general.” In addition, many teachers commented that a great deal of time was spent on training with the new evaluation tools, with another teacher commenting, “There was a real effort made to make sure we all really understood it.” Many also expressed, however, that the instrument was overwhelming and cumbersome, with one teacher commenting that “they were a little bit difficult to navigate especially at first,” and another commenting “I feel like to really understand the rubric you have to be very familiar with it, stare at it and move the papers around constantly, memorize it before you can understand it, it was really overwhelming.” Another shared, “I think the other thing that is really tough is, as an educator, I want to be the ‘excellent,’ right? I want to be the ‘highest one,’ and the fact that, that’s pretty much impossible to achieve is disheartening.” Despite these reservations, it is apparent that the rubric was perceived as representing a high standard of effective teaching.
**Summary of results to research question two.** Overall, these results were more positive than in response to those items considered in response to Research Question 1, but there remained significant differences between the middle school and high school participants. Teachers reported that the self-reflection process using the teacher rubric caused them to consider their own practice (64% agreed or strongly agreed). However, the difference between the high school (45%) and middle school (85%) as well as District A (85%) and District B (44%) continued the existing trend. Overall, the greater majority of the population (83%) agreed or strongly agreed that the rubric was a comprehensive representation of the traits of effective teaching. Interestingly, even amongst those groups typically scoring lower on earlier survey questions, they too looked favorably upon the rubric as a comprehensive representation of the traits. Seventy-four percent of District B participants and sixty-nine percent of high school participants agreed or strongly agree. And 95% of District A participants agreed.

**Overarching Themes related to Teachers’ Perceptions of the Value and Impact of the New Evaluation System**

The last three questions of the survey were open-ended and gave participants the opportunity to tell the researcher more of their experience with and perceptions of the new evaluation instrument. Participants responses were categorized and then synthesized from across responses and are reported below.

**Overall, what are your thoughts on the new teacher evaluation system?** The first open-ended question asked teachers for their overall thoughts on the new teacher evaluation system. This general question elicited negative responses on 50% of the responses; participants shared comments centered on issues of complexity of the system, the overwhelming process, and citing time taken away from preparing for instruction. For example, one teacher commented
that, “The process is burdensome and takes away from time which I would use to grade assignments and prepare new materials to use in class.” Positive responses were found in only 11% of the participants with overriding themes including the rubric being a good indicator, that the system helps the self-reflection process, and that it is fair and accurate. One respondent shared, “I think the system helps teachers be more self-reflective on their practices and possibly causes them to alter or improve lessons and approaches to instruction.” The remaining 26% either expressed neutral opinions or mixed opinions citing positive and negative aspects of the program.

**What do you like about the new teacher evaluation system?** The second question, what do you like about the new teacher evaluation system, produced positive replies in 87% of the 46 participants. Overall, teachers liked the opportunity to provide their own evidence, a clear standard for evaluating teacher performance, and the reflective nature of the system. Furthermore, teachers expressed positive perspectives on the opportunity to meet and discuss their progress with their evaluator, the opportunity to focus on their practice, and the opportunity to keep teachers current in their field. The 13% who responded negatively to the new process shared that it was too cumbersome and complex as well as concern over student surveys eventually becoming part of the process.

**What do you not like about it?** The third open response question, “What do you not like about the evaluation system?”, again prompted for deeper insight into teacher perceptions. Obviously, due to the nature of the question, the responses were overwhelmingly negative. However, in many of the responses teachers qualified their concerns. The primary concern, represented in 54% of the responses, was that the new evaluation system was too time
consuming and this process takes time away from working with students and other essential teaching tasks. As one teacher shared,

What I like the least about it is that, in the words of my colleague, it’s "one more thing we have to do." Because so many of the indicators are not things that can be seen during classroom observations, the onus is on the teacher to provide that proof. This is not the negativity of a beleaguered veteran speaking -- this is reality. Each year there are more professional responsibilities given to us, and rarely is anything taken off our plate. I think that that time might not be better spent either preparing further for classes or assessing students.

Another teacher shared similar thoughts,

I do not like that it creates more paperwork and processes for me. There is a finite amount of time in a day and ultimately in a school year. I would rather be creating dynamic lessons or assessing student learning and making thoughtful decisions about my teaching practice rather than spending time creating evidence that I am doing this.

The second major concern expressed in this question (9%) was the accuracy of the instrument and data collection process, as one teacher shared, “It does not and cannot accurately say whether a professional is doing a good job or not.”

**Summary of the overarching themes of the open-ended survey questions of the perceived value and impact of the new teacher evaluation system.** When asked generally for their impressions, 50% of the respondents replied negatively and shared concerns over the complexity of the new system as well as the time it takes away from students. In contrast when asked what they liked about the new evaluation system, teachers shared they liked the opportunity to provide evidence, the clear expectations and standards for evaluation, as well as
the opportunity to meet and discuss their performance with their evaluator. Those that offered a negative opinion consistently cited the complexity of the evaluation tools and several were overwhelmed by the rubric specifically. On the other hand, participants liked the standard expectations although they often cited they were difficult to understand at first.

**Research Question #3: Do teachers perceive use of the tool and engagement in the process as an effective means toward changing their teaching practice for greater student achievement.**

The next question reflects on the research goal of examining teachers’ perceptions in their use of the new evaluation tools to change their teaching practice to increase student achievement. Ultimately, the goal of the new evaluation system is to foster improved teacher performance (pedagogy) and improved student achievement (Morelock, 2008; Massachusetts Department of Elementary and Secondary Education, 2011).
Table 8

Setting SMART goals has helped to focus my efforts on increasing student achievement.

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither</th>
<th>Agree</th>
<th>Strongly Agree</th>
<th>Rating Average</th>
<th>St. Dev.</th>
<th>Response Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5 Years</td>
<td>0%</td>
<td>0%</td>
<td>20%</td>
<td>80%</td>
<td>0%</td>
<td>3.80</td>
<td>0.45</td>
<td>5</td>
</tr>
<tr>
<td>6-10 Years</td>
<td>6%</td>
<td>29%</td>
<td>24%</td>
<td>41%</td>
<td>0%</td>
<td>3.00</td>
<td>1.00</td>
<td>17</td>
</tr>
<tr>
<td>11-15 Years</td>
<td>4%</td>
<td>23%</td>
<td>19%</td>
<td>50%</td>
<td>4%</td>
<td>3.27</td>
<td>1.00</td>
<td>9</td>
</tr>
<tr>
<td>&gt;15 Years</td>
<td>5%</td>
<td>10%</td>
<td>24%</td>
<td>52%</td>
<td>10%</td>
<td>3.52</td>
<td>0.98</td>
<td>21</td>
</tr>
<tr>
<td>Elementary School</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>100%</td>
<td>5.00</td>
<td>0.00</td>
<td>2</td>
</tr>
<tr>
<td>Middle School</td>
<td>0%</td>
<td>5%</td>
<td>14%</td>
<td>76%</td>
<td>5%</td>
<td>3.81+</td>
<td>0.60</td>
<td>21</td>
</tr>
<tr>
<td>High School</td>
<td>7%</td>
<td>24%</td>
<td>28%</td>
<td>41%</td>
<td>0%</td>
<td>3.03+</td>
<td>0.98</td>
<td>29</td>
</tr>
<tr>
<td>District A</td>
<td>0%</td>
<td>4%</td>
<td>20%</td>
<td>72%</td>
<td>4%</td>
<td>3.76^</td>
<td>0.60</td>
<td>25</td>
</tr>
<tr>
<td>District B</td>
<td>9%</td>
<td>30%</td>
<td>22%</td>
<td>39%</td>
<td>0%</td>
<td>2.91^</td>
<td>0.60</td>
<td>23</td>
</tr>
<tr>
<td>District C</td>
<td>0%</td>
<td>0%</td>
<td>25%</td>
<td>25%</td>
<td>50%</td>
<td>4.25</td>
<td>0.96</td>
<td>4</td>
</tr>
<tr>
<td>Dist. A- M.S.</td>
<td>0%</td>
<td>5%</td>
<td>10%</td>
<td>80%</td>
<td>5%</td>
<td>3.85#</td>
<td>0.59</td>
<td>20</td>
</tr>
<tr>
<td>Dist. B - H.S.</td>
<td>9%</td>
<td>30%</td>
<td>22%</td>
<td>39%</td>
<td>0%</td>
<td>2.91#</td>
<td>1.04</td>
<td>23</td>
</tr>
<tr>
<td>Non-Dist. B-H.S.</td>
<td>0%</td>
<td>0%</td>
<td>50%</td>
<td>50%</td>
<td>0%</td>
<td>3.50</td>
<td>0.55</td>
<td>6</td>
</tr>
<tr>
<td>OVERALL</td>
<td>4%</td>
<td>15%</td>
<td>21%</td>
<td>54%</td>
<td>6%</td>
<td>3.42</td>
<td>0.96</td>
<td>52</td>
</tr>
</tbody>
</table>

^ Teachers from District A were overall in much more agreement with the survey statement than teachers from District B (Fisher Exact Test p=0.010)
+ Middle school Teachers were significantly more in agreement than high school teachers in (Fisher Exact Test p=0.005)
#District A middle school teachers significantly responded more positively than District B high school teachers (Fisher Exact Test p=0.002)

Table 8 demonstrates that teachers overall believed that setting SMART goals helped teachers focus their efforts on increasing student achievement, with an overall response average of 3.42 (σ = 0.96). However, agreement to his statement was not very high. Only 60% overall
reported they either agreed or strongly agreed with the statement. Subgroups that agreed the most were those with more than 15 years of teaching experience (62%), middle school teachers (81%), District A teachers (76%), and, in keeping with positive responses to other items throughout the survey, the district A middle school teachers (90%). Subgroups with the lowest agreement responses included 6-10 years of teaching group (41%), high school teachers (41%), district B teachers (39%), and the district B high school teachers (39%).

Once again significance was found between District A and District B teachers (Fisher Exact Test, \( p=0.010 \)), the middle and high school teachers (Fisher Exact Test, \( p=0.005 \)), and the District A middle school and District B high school (Fisher Exact Test, \( p=0.002 \)). The teaching experience groups did not show significant differences using the Fisher Exact Test. The trend exhibited once again in this data draws one to the differences between District A and District B, the middle School and high school participants, and perhaps more precisely between District A middle school and District B high school.

The focus group responses were supportive of the feeling that the reflection and goal setting process was positive and had the potential to influence positively student achievement. As one participant shared,

I think having a set way to have teachers reflect on what they are doing, I think that that is something we don’t normally make time for, I think that has to improve instruction and if you improve instruction it’s going to improve on student learning.

Another teacher added, “If you are constantly reflecting what you doing within the classroom, it’s going to make for a better teacher.” A third teacher stated, “I think it would, because if I'm being a better teacher, if I'm being more aware and conscious, then I would hope that that would improve my students' achievement.”
One teacher in a focus group did not agree, and when asked if they felt the evaluation system has the capacity to improve student achievement, he succinctly responded, “NO.”

**Summary of the results of research question three.** Overall, the majority of participants agreed or strongly agreed that the setting of SMART goals enabled the focusing of efforts on increasing student achievement (60%). As increasing student achievement is a critical and significant purpose of the new evaluation system, it is significant that the majority of participants agree that the SMART goals help them to focus their efforts on increasing student achievement. At the same time, it was 54% that agreed and only 6% that strongly agreed. However, it is also important to highlight that only 20% disagreed or strongly disagreed with the statement.

As for responses across the various groups, and consistent with the response to other survey items, middle school participants were in much greater agreement (81% agree or strongly agree) compared to the high school participants (41% agree and 0% strongly agree), and District A participant responses were higher (76% agree or strongly agree) than District B participant responses (only 39% agree or strongly disagree).

In the focus groups, participants shared that the process helped them to focus on student achievement and the process should make for better teaching.

**Summary of Findings**

The findings of this study indicate that overall participants reported that they agree with key goals of the evaluation system to improve instructional practice and student achievement. Overall, the majority of participants agreed with the statement that the evaluation process caused positive changes in their teaching practice, with 50% agreeing and 6% strongly agreeing with that statement. However, when asked if the evaluation process resulted in changes in their
teaching practice the results were lower, with an average response of 2.92, with only 39% agreeing or strongly agreeing with this statement and 43% disagreeing or strongly disagreeing. Interestingly the middle school group reported the highest results to these two questions with a 3.81 average and 81% agree and strongly agree in regards to evaluation system causing positive changes in their teaching practice and 3.43 and 57% agreeing or strongly agreeing with the statement the evaluation process resulted in changes in their teaching practice. In contrast, District B was the low responding group (2.82 and 34% agree or strongly agree) to the question regarding positive changes in one’s teaching practice as a result of the evaluation system. Whereas the 11-15 experience group was the low responder (2.92 and 39% agree or strongly agree) to the statement that the evaluation process resulted in changes in the individuals teaching practice.

When looking at the self-reflection process and use of SMART goals a majority reported, they found value in the process. When asked if self-reflection caused the respondents to consider their teaching practice 67% agreed or strongly agreed with the statement and overall the average was 3.54 on the 5-point Likert scale. Overall, the response to whether SMART goals helped teachers to focus on student achievement resulted in a 3.42 average and 60% agreeing or strongly agreeing with the statement. The last inquiry related to SMART goals asked teachers to respond to whether the use of SMART goals improved teaching practice. This produced a 2.97 average in the 5-point Likert scale and 56% agreeing or strongly agreeing with the statement lower that the other questions in this category. It is interesting to note that district B was the low reporting group in each of these three questions; self-reflection causing consideration of teaching practice 3.17 and 44%, SMART goals helping to focus on student achievement 2.91 and 39%, and lastly SMART goals improving their teaching practice 2.83 and 30%. Meanwhile district A was the
highest reporting group or a close second; self-reflection causing consideration of teaching practice 3.95 and 75%, SMART goals helping to focus on student achievement 3.85 and 85%, and lastly SMART goals improving their teaching practice 3.80 and 75%.

In regards to the implementation of the new evaluation system the participants felt strongly that the rubric represented the traits of effective teaching with a 3.69 average and 83% either agreeing or strongly agreeing with statement. Another strong response was obtained to the statement that their evaluator was able to gather enough evidence to assess them accurately with a 3.73 average and 75% either agreeing or strongly agreeing. Even those groups that were less positive were still either above or slightly below 3.0 on the Likert scale. The 1-5 year experience group was the lowest scoring group on the rubric represented the traits of effective teaching with a 3.40 and 60% agreeing and 0% strongly agreeing with the statement. When responding to the statement that their evaluator was able to gather enough evidence to assess them accurately the 6-10 year experience group averaged 3.41 and 53% reporting that they either agreed or strongly agreed. District A responded highest to the statement that the rubric represented the traits of effective teaching with a 4.0 on a five point Likert scale and 95% reporting that they agree or strongly agree with the statement. The middle school group responded highest to the statement that their evaluator was able to gather enough evidence to assess them accurately with a 3.86 average and 85% either agreeing or strongly agreeing.

When the survey asked, what the teachers overall thoughts were on the evaluation system 50% shared a negative opinion centered on the complexity of the system or that it was time consuming and negatively influenced their ability to plan and work with students. Reponses to the question of what they liked about the system resulted in 87% positive responses centered around the themes of the ability to provide evidence, clear standards for teaching, and the
reflective nature of the evaluation process. In the last question, what do you not like about the evaluation system, 54% cited issues with the complexity, amount of paperwork, and time taken away from students and other important tasks in teaching.

It is important to note that consistently significant differences appeared between District A and B, middle and high school, as well as between District A middle school and District B high school. These differences cause one to ponder whether these differences are between districts or middle schools and high schools, and will be discussed more fully in the Discussion of Findings in Chapter 5.

In conclusion, many teachers reported that parts of the new evaluation system were helpful and assisted in the improvement of their teaching practice, but many teachers also disagreed with the statements regarding the desired outcomes of the system. Many participants responded that the self-reflection and SMART goal aspect of the process enabled teachers to reflect on their teaching practice and set goals for improvement both in practice and student achievement, but there were also a good number (as many as half in some cases) that disagreed. Furthermore, the rubric was reported to be a good indicator of effective teaching practice and evaluators were able to collect sufficient evidence to evaluate accurately.

Chapter V: Discussion of the Research Findings

This chapter will provide a summary and discussion of the findings as well as commentary on the significance of this study. The chapter will begin with a summary of the problem of practice, followed by a review of the methodology, a summary of the findings, a discussion of the findings in relation to the theoretical framework, a discussion of the findings in relation to the literature review, a review of the validity and limitations of the study, and an examination of the next steps in research.
Summary of the Problem of Practice

As mentioned earlier in this paper, the goal of the new Massachusetts evaluation system is to provide for the improvement of teaching practices and ensure that only highly effective individuals lead instruction in our classrooms (Rothstein, 2010; Massachusetts Department of Elementary and Secondary Education, 2011). The new Massachusetts model for educator evaluation was based heavily on Danielson’s model (2007), which promotes the use of a common framework to guide professional conversations on effective teaching. The new Massachusetts evaluation model represents a significant shift in the evaluation process for a majority of locally controlled districts. The intent of this study was to examine teachers’ perspectives of the evaluation process in its pilot form across three different districts, guided by the following research questions:

1. Do teachers perceive the new Massachusetts teacher evaluation system as a valid and useful tool for changing their teaching practice?
2. Do teachers perceive the evaluation process of self-reflection, goal setting, and evaluation using the rubric as a valid and worthwhile assessment of effective teaching?
3. Do teachers perceive their use of the tool and engagement in the process as an effective means toward changing their teaching practice for greater student achievement?

Summary of the Findings

According to the results of this study, one would consider teachers’ perspectives of the utility of the piloted teacher evaluation were rather mixed. With this said, it is important to note that this pilot was not a statewide-standardized effort. The state of Massachusetts asked districts to “early adopt” the prototype evaluation instrument and participate in discussions on its implementation. These eleven early adopter districts field-tested the evaluation tool in different
ways, some districts engaging a small number of teachers and others employing a large district effort. Below the researcher provides synthesis and explanations for the groups responses and overall findings.

There were significant differences across groups participating in the study. It is important to note that some groups looked much more favorably upon the potential benefits of the new evaluation system than some other groups. In particular, participating high school teachers were much less likely to agree positively to a number of the survey items than middle school teachers. In particular, the district A middle school teachers were far more in agreement across survey items than the district B high school teachers.

In an attempt to explain this potential difference in perceptions, the researcher recognized that there were wide variations across groups in previous teacher evaluation systems as well as the amount of people piloting the new evaluation system. District A represented the majority of middle school teachers and is a district that previously used walkthrough observations and goal setting similar to the processes involved in the new pilot. This was highlighted by one teacher sharing, “We were already ahead of the game, we had already had walk-throughs, already did goal setting and Smart Goals, We already had half year reflections.” In contrast, District B had only previously used a traditional evaluation method of pre and post meetings and scheduled observations. With this as a backdrop, a teacher from District B put it this way, “It’s much more intrusive, and, and certainly it’s kind of reflecting more of the stuff you did anyway. But it’s just an enormous change.” District C was a district that used a very small sample size of four teachers, who volunteered for this pilot and worked in three different buildings. This small size might be an important factor in teacher perceptions. As one teacher from this district noted, “I like the feedback, that my principal gives us right away.”
An interesting pattern emerged where District A, District A middle school, and the middle school groups scored higher in most categories and significant differences were found between them and the District B, District B high school, and high school groups. These differences were highlighted in the research findings. It is important to reiterate that District A had previously incorporated several of the new methods now used in the new evaluation process, whereas this process for District B was an entirely new enterprise. This, combined with the fact that District A participants also were the majority of the middle school group, and that District B represented a majority of the high school group may explain some of the consistency in the results and comparisons between these two groups. Unfortunately, due to the limited data available it was impossible to determine if these differences were more of a difference between the districts or the differences between high school and middle school teachers. This is an area that should be explored in future research.

Another interesting pattern was found with those who had been teaching for six to ten years. Although this group did not consistently differentiate itself from other specific teaching experience groups, it did distinguish itself consistently. This group scored lower in those areas where the evaluation process and its effect on their teaching practice was assessed, but high in the area on the ability of SMART goals to improve student achievement. From the focus group discussions, it was apparent that this group had undergone a similar reflective portfolio process in their pursuit of master’s degrees. As one teacher shared, “I had to do that for my undergrad and graduate work.” As these previous efforts had already resulted in professional reflection and change it is probable that this group did not see much effect on their practice at this point in their careers.
Teacher perceptions of the new Massachusetts teacher evaluation system as a valid and useful tool for changing their teaching practice was mixed and varied across groups. This study showed that overall, the results were mixed, and perhaps dependent on the context in which the evaluation system was employed. Overall, a slight majority of teachers agreed or strongly agreed (56%) that smart goals helped them focus on improving their practice. However, 23% disagreed or strongly disagreed, and 21% reported that they neither agreed nor disagreed. The pattern of differences between the middle school and high school, Districts A and B, and District A middle school and District B high school are very pronounced in these questions results. The difference between the responses of the District A middle school teachers (75% agreed or strongly agreed) and the District B high school (30% agreed or strongly agreed) was most pronounced. This raises the question if these differences are representative of the structural differences between high schools and middle schools or representative of the amount relative change for each district with the new system from previous experiences in teacher evaluation systems.

Regarding the survey question having to do with teachers’ responses to the question of how the evaluation process caused positive changes in their teaching practice, the overall average response was slightly less than the neutral 3.0, with only 39% agreeing or strongly agreeing with the statement and 43% disagreeing or strongly disagreeing. In keeping with the noted trends between middle and high schools and districts, 57% of middle school teachers agreed or strongly agreed while 59% of high school teachers disagreed or strongly disagreed. This pattern is also reflected in the District A and District B comparison. In District A, 55% agreed or strongly agreed while 56% disagreed or strongly disagreed in District B. As for the difference between the District A middle school and District B high school, the 55% of the District A middle school
teachers agreed while 56% of the District B high school disagreed or strongly disagreed. One should keep in mind that this question was asked after only one year using the new system and although it elicited strong reactions in either direction, it may be too soon to expect participants to assess truly their progress.

The next question collected the most positive overall response and asked if evaluators were able to collect enough evidence to evaluate them accurately. Seventy-five percent shared that they either agreed or strongly agreed with the question and the average was well above the neutral (3.73). Although there were some differences across groups, it is important to note that no group had larger than a 20% in the disagree, strongly disagree categories. This demonstrates that teachers felt confident in their evaluators’ ability to collect enough evidence to evaluate them accurately. One should note that these pilots represented less than the full evaluation loads experienced by administrators presently. This might be an important question to ask under the present situation of full implementation of the new evaluation system.

The last question, related to the first research question asked whether the new evaluation process changed their teaching practice and elicited a mixed bag of results, perhaps indicating that it was too soon to tell if this important goal of this initiative was reached. Although overall 56% agreed or strongly agreed that the new evaluation model positively changed their teaching practice 27% disagreed or strongly disagreed. The different perceptions once again are evident in the subgroups where 80% of District A middle school agree or strongly agree compared to 34% of District B high school, 68% of District A agree or strongly agree compared to 35% of District B, and 81% of middle school teachers and 34% of high school teachers either agree or strongly agree. Although the results are varied and it may be too soon to assess accurately this
particular issue, the question of whether this process is better received in a particular district or type of school continues.

**Teacher perceptions of the evaluation process, including goal setting and using the rubric as an instrument to assess effective teaching, were also mixed across groups.** Again the same groups demonstrate lower agreement rates in this survey question regarding the value of the self-reflection process and consistent with previous trends the high school (45%), District B (44%) and District B high school (44%) all showed lower percentages for the agree and strongly agree categories. One teacher shared, “In years past my goals have always been outcome related, and this year I looked more internally which was very different that anything I had ever done in the past.” It appeared from the results that although the process was difficult, teachers found value in the process,

But when you create that self-assessment you kind of take a step back and you reassess yourself and ask why am I teaching like this and why don’t I try something different so you know you kind of step back from the routine of it all and try to change your teaching style.

Another volunteered that this was a difficult process, “I do agree that the process was very difficult even if they made it clear as they thought they made it and with as much direction as we did get.”

When asked whether the Massachusetts Teacher Performance Rubric was a comprehensive representation of effective teaching every group responded positively with a score of 60% or greater in the agree strongly agree categories except the non-district B high school teachers (33%). One teacher explained her thinking on the rubric, “I think they’re good goals. They’re good standards in general.” However, a few found them overwhelming, “I think
that some of these (are) a little bit redundant and I think that some of them are very difficult to prove (with) pieces of paper.” The open-ended survey questions and the focus group interviews highlighted that this process was seen as time consuming, and that teachers feared this was time taken away from student-centered activities and would negatively impacting students.

**Teachers perception of the use of the tool and engagement in the process as an effective means toward changing their teaching practice for greater student achievement also demonstrated mixed results.** In the third and arguably most important research question was addressed in focus group discussions and in a survey question if setting SMART goals help one to focus their efforts on increasing student achievement as shown in table 8. The survey results were mixed and lower than the focus group question were 90% responded positively. The focus group question, overall do you feel the evaluation system has the capacity to improve student achievement if so how, received only one negative response (10%) out of 10 responses to this question. The majority of the responses were similar to this,” I do, I think having a set way to have teachers reflect on what they are doing, that is something we don’t normally make time for, I think that has improve instruction and if you improve instruction it’s going to improve student learning.” The one negative response was a simple no, after a previous statement that he did not feel his evaluation was accurate. This discrepancy might be due to the fact that seven out of the ten teachers who participated in focus group discussions were from District A. Another explanation might be that seven out of the ten focus group participants were from middle and elementary schools. In table 8 overall the results indicate that 60% of the teachers surveyed agreed or strongly agreed that setting SMART goals focused their efforts on increasing student achievement. However, we still see those groups with lower agree and strongly agree percentages, high school (41%), District B (39%), and District B high school (39%). The trend
continues with the opposite groups having significant higher agree and strongly agree scores middle school (81%), District A (76%), and District A middle school (85%). As a key goal of this initiative is to provide an instrument and self-reflective process to improve instruction and student achievement, these results indicate that teachers perceive this evaluation tool as possessing that potential however the results are only moderately positive in the use of SMART goals.

**Discussion of the Findings in Relation to the Theoretical Framework**

The finding of this study relates directly to Performance Improvement Theory (PI), where one is focused on clarifying the goals and behaviors as well as harmonizing the goals and behaviors of individual contributors, work process owners, and organization leaders (Swanson, 1999). The new Massachusetts evaluation model includes a detailed teacher performance rubric that characteristics of effective teaching and measurable SMART goals focused on both their professional practice and improving student achievement (Massachusetts Department of Elementary and Secondary Education, 2011). The SMART goal process of developing measurable goals with one’s evaluator, in the areas of professional practice and student achievement, is a key aspect of the new evaluation process. The study participants consistently indicated that the self-reflection and SMART goal process was moderately valuable as evident in tables 2, 6, and 8. In these examples, a majority of the respondents shared that SMART goals helped them focus on improving their teaching practice (56%), that self-reflecting caused them to reflect on their teaching practice (64%), and that they believed the use of SMART goals would result in an increase in student achievement (60%). Overall 73% of survey respondents (table 7) agreed or strongly agreed that the Massachusetts Teacher performance Rubric was a comprehensive representation of effective teaching which aligned with the tenor of performance
improvement theory where goals are clarified and focused. The clarity of the Massachusetts Teacher performance Rubric enables both those being evaluated and those doing the evaluations to focus on key behaviors and characteristics, provide training in specific areas, thus harmonizing the efforts of both sides to meet these performance standards.

These findings also align with Schaffer and Keller’s (2003) assertion that stakeholders should be involved in the design of performance evaluation systems early to ensure success. The process of developing SMART goals is a collaborative effort between the person being evaluated and the evaluator. This process enables stakeholders a voice in designing this aspect of their evaluation instrument, causes individuals to reflect on their practice and their students’ achievement, and create measurable goals to address and foster improvement. However, this was an entirely new and somewhat overwhelming process for some, as one teacher shared, “Complex, hard to understand first time you do it”, and “While goal-setting is admirable and creates focus, I find it very difficult to manage creating something new during the school year.” While the new process does provide the opportunity to be involved with designing aspects of the evaluation consistent with Schaffer and Keller (2003), work is still needed in the harmonizing goals aspect as described by Swanson (1999).

In regards to having clear behavioral expectations, the Massachusetts Teacher Performance Rubric clearly delineates the expectations for effective teaching in the state of Massachusetts (Massachusetts Department of Elementary and Secondary Education, 2011). This is especially evident in Table 7, where 83% of participants shared that they felt the rubric illustrated the aspects of effective teaching, and in Table 5 where 56% agreed that the evaluation process positively changed their teaching practice. Although only 39% responded they agreed or strongly agreed with the assertion that the evaluation process changed their teaching practice.
(Table 3), this may be because the opportunities to receive feedback, reflect, and revise goals were in their beginning stages and not completely developed as a process or product. One teacher supported this assumption, “I think the system helps teachers be more self-reflective on their practices and possibly causes them to alter or improve lessons and approaches to instruction.” Another offered, “I think setting specific goals can be helpful as it pertains to teaching methods. My teaching methods are always evolving due to professional development, reading, etc. and I don't think the evaluation process resulted in any significant changes in how I teach.” Another teacher illustrated the range of thoughts on the new system,

It's too process oriented. Too many steps involved which takes too much time away from what we could be doing for our classes. I like the goal setting process, which was similar but more focused than our previous system. There was not nearly as much emphasis on the attention we give our goals throughout the year. Now we have more accountability.

It is apparent that although the goals of the system are clearly defined, the harmonization of these goals will take more time and effort. However, to be fair this study evaluated a pilot, and in the second year of implementation of the New Massachusetts Educator Evaluation system, the Massachusetts Department of Elementary and Secondary Education has addressed these issues by developed training modules.

Based on these results, it is apparent that the fundamental aspects of Performance Improvement theory, clarifying goals and behaviors, and harmonizing the goals and behaviors all stakeholders are addressed in the new evaluation model.

**Discussion of the Findings in Relation to the Literature Review**

The results of this study elucidate three areas covered in the Literature review. These three primary areas were: (1) definitions of effective teaching and how effective teaching should
be measured; (2) definitions and examples of effective evaluation models; and (3) research on the impact of evaluation models on teachers’ instructional practice.

**Definitions of effective teaching and how effective teaching should be measured.**

According to the Gates Foundation (2009), effective teachers work in ways that result in student gains in student achievement measures. Of the participating teachers, 64% agreed or strongly agreed that the self-reflection and goal setting processes and use of the Massachusetts Teacher Performance Rubric was a comprehensive guide representation of effective teaching practice. Furthermore, 73% either agreed or strongly agreed with the statement that the rubric represented the standards and indicators of effective teaching. This was also supported through focus group and open response offerings such as “They’re good standards in general” and “I think it is more accurate way to evaluate teachers.” Many of the sources cited in the literature review described specific factors that they define as essential in teaching including management of the classroom, clear goals, the ability to judge work and give constructive feedback, caring for their students, strong instructional methods, and the willingness to improve those instructional practices (Berry, Daughtrey, and Wieder, 2010; Darling-Hammond, 2009; Giovannelli, 2003; Mortimore, 1994; Porter and Brophy, 1998). However, it seems that the rubric, based on Danielson’s (2007) work, resonated with the teachers and may have captured many of these skills indicated by the other authors.

**Definitions and examples of effective evaluation models.** The U.S. Department of Education’s Race to the Top initiative required that states include student achievement and annual teacher evaluations in their waiver applications (U.S. Department of Education, 2009). In this effort, the state of Massachusetts developed a system incorporating observations related to a rubric of teaching qualities based on the Danielson model (2007). This rubric incorporates many
of the qualities of effective teaching summarized in the literature review and approaches the process through multiple angles (Darling-Hammond, 1999; Kane, 2012; Morelock, 2008). Participating teachers’ responses to Table 7 indicates that they view the rubric as representative of effective teaching practice (73% agreed or strongly agreed in Table 7). In addition, they valued the self-reflection on one’s practice as engaged in the evaluation process (64% agree or strongly agreed in Table 6). Although this study investigated the perceived impact of an initial pilot of the new evaluation process, 56% reported it had a positive effect on their teaching practice in a relatively short amount of time (Table 5). In the open response questions one teacher shared, “I think the system helps teachers be more self-reflective on their practices and possibly causes them to alter or improve lessons and approaches to instruction.” Another offered, “The indicators are pretty comprehensive, covering most of what good teaching means.” However, another illustrated pessimism when asked about the new evaluation system, “Very complex, untested and untried at this point.” These results are consistent with the research in both the definition and the perception of this model as an example of an effective evaluation model.

**The impact of evaluation models on teachers’ instructional practice.** As mentioned above in table 5, overall 56% of the participants reported that the new Massachusetts teacher evaluation process had a positive effect on their teaching practice. This is in conflict with Adams (2009) and Huckstadt (2011) who either found little or no teacher reporting of improving teachers’ instructional practice were here a slight majority reported the new evaluation implemented in Massachusetts did improve their professional practice. It is important to point out that the “six-to-ten years of teaching” group might be representative of individuals who had previously used a portfolio and standards processes in their degree programs, which harbor
similarities to the new evaluation process. This may be a possible explanation for this group’s lower reporting in this category (41% agree 0% strongly agree in Table 5). As District B participation consisted only of a high school and comprised the majority of the high school group overall, this may explain the similarity of the low level of agreement in this group. Furthermore, the new evaluation system was drastically different from the previous evaluation instrument in District B, and many expressed that they were either unprepared or overwhelmed with the process. One teacher from District B shared, “The amount of time I spent worrying about it (the evaluation process) and learning how to do it and what went into it. That was just enormous.” This may be a possible explanation for this group’s low agreement scores. Although previous efforts in merit pay or pay-for-performance have failed or been a mixed bag (Burns, et.al, 2009; Glazerman & Seifullah, 2010; Springer, et.al, 2010) it is premature to evaluate this system’s impact on student achievement results, as the links to student growth are still in the process of being developed. However, early indications of its impact on professional teaching practice are encouraging, as one teacher shared, “I think the system helps teachers be more self-reflective on their practices and possibly causes them to alter or improve lessons and approaches to instruction” and one would hope that there is the possibility of improved future student achievement.

**Validity and Limitations**

Potential issues of validity of this study were addressed in chapter three and included objectivity as well as internal and external efforts to ensure validity. In order to ensure that the biases of the researcher were not influencing the collection of data in this study, the same survey was used in all locations and the focus groups were conducted with the same questions and in the same method and order. Furthermore, focus groups were conducted outside of the researcher’s
district and in neutral locations. Throughout the entire process, the researcher bracketed his personal feelings and biases so as not to influence participants’ responses. To ensure validity of the focus group data, the transcripts of the discussions were reviewed with focus group participants to ensure accuracy. The researcher did not meet or contact survey participants unless they volunteered for and participated in the focus groups. Furthermore, to ensure the validity of the responses the identities of those that participated in the survey and focus groups were not shared with the school and district administrators.

Limitations of this study include primarily the small number of participants in this study. One needs to understand that only eleven districts statewide, participated in this pilot. These eleven “early adopter” districts had different levels of implementation of the new evaluation system with two reporting to this researcher, when approached, that they never moved beyond discussions with their local union. Furthermore, in some districts participants in the pilot were limited, again reducing possible participants in this study. In two other districts, this researcher's request for participation was declined for various reasons. This factor of limited numbers was also exacerbated by the fact that in two of the districts with larger numbers of participants, this researcher was limited by local administration as to where this research could be conducted. Although the participant numbers were small, so was the original pilot, and this study was only designed to provide initial indications of the implementation of the new evaluation system.

**Significance of the Study in the Field**

The rationale behind the new teacher evaluation systems is to ensure that high quality teaching is occurring in our classrooms (Rothstein, 2010, Massachusetts Department of Elementary and secondary Education, 2011). This study examined teacher perceptions of the pilot of the Massachusetts new evaluation system in regards to the process and impact on teacher
practice. It is important to note that the most positive responses originated in the district that had previously used similar components to those in the new evaluation system. This district also exhibited a collaborative culture and a high level of trust when piloting the evaluation system. Overall teachers reported positively with the process, tools, and impact of the new system. It is obvious, that the focus on goals and the observable characteristics of effective teaching are perceived as having a positive impact on professional practice. Furthermore, participants report that they believe this process will have a positive effect on student achievement. However, they also raised serious concerns over the impact of the time spent on these tasks and particularly the impact it will have on the core responsibilities of teaching and their ability to attend to student needs. The teachers perceptions are in agreement with performance improvement theory, in that clear goals and focused feedback are essential in improving performance (Swanson, 1999). Moving forward, the potential is present in the new system to improve teacher performance when the evaluator and teacher work together with clarity in this process.

**Final Words**

In conclusion, while this study has supported the potential for the improvement of teaching practice as reported by teachers through the new Massachusetts teacher evaluation process, the initial perceived utility of the teacher evaluation system vary widely, dependent on the school if not the district. These demonstrated variations appeared most consistently between schools, high school and middle school, and districts, District A and District B. It is important to remember that District A composed most of the middle school category and District B composed most of the high school category. This fact makes it uncertain if these results are due to District A having previously used a more progressive evaluation system than District B, or if these results demonstrate a significant differences in the perceptions of high school and middle school
staff. To further complicate matters, the non-District B high school category had mixed results sometimes in opposition to District B high school and other times agreeing. In the focus group discussions it was obvious that District A participants were more comfortable in the evaluation process citing several times that they had engaged in some of these processes before, where District B participants seemed to be more overwhelmed by the scope of the new procedures. However, the majority of the participants in District A were middle school staff and District B consisted entirely of high school teachers, further augmenting the debate if this difference was based on the district or level of school.

It is important to note that these pilots were small in scope and application, the smallest of which was District C, a large urban district, which only had four teachers pilot the new system. This evaluation process, although worthwhile, requires a significant increase in the workload of both the evaluator and person being evaluated. One teacher summarized this thought,

The ideas and standards are correct in that they do contribute to being a good teacher, but the whole system and evaluation plan is unrealistic to do well. In order for an evaluator to see that a teacher is doing all this well they need to be observing this teacher a lot and being familiar with all that they do. It's an incredible amount of work on the administrator and teacher's sides to put together good evaluations and documentation proof of it all and it's just not realistic that we can do this in a regular school year while teaching.

It will be interesting to see if the quality and clarity of feedback and goals can be maintained once every teacher is placed on the evaluation cycle, greatly increasing evaluators’ workload.
Next Steps

This study represents a first look at the Massachusetts teacher evaluation system as proposed and piloted in 2011-2012. At this time, all districts in Massachusetts were in either their first or second year of implementing the new system. As such, the potential for a larger sample size is present to add validity to this study. Furthermore, although there are many factors involved in student achievement outcomes, and the data systems to measure student achievement have their flaws, it would be prudent to examine if there is a correlation between teacher ratings and student achievement.

The potential also remains for continued research into this topic to examine differences in individual district implementation strategies and their perceived and real success in changing teacher practice. Each district was left to negotiate individually with their local unions the specific implementation of the new evaluation model. Although the department published new regulations governing a new evaluation process, essentially scaffolding basic requirements, areas were open for local discussion (Massachusetts Department of Elementary and Secondary Education, 2011). Although there is more similar than different, some slight differences exist between different districts here in Massachusetts. Research examining different implementations across several districts in the form of case studies might discern both successful strategies as well as challenges to successful implementation in some districts. Such information could potential inform how districts and schools could more successfully implement the evaluation system, strengthening its effective use across the state, and possibly resulting in greater student achievement.

Another potential area of study would be to solicit the opinions and perceptions of evaluators, administrators, working in the new evaluation system. This study only solicited
teacher perceptions, and value could be found in ascertaining the evaluator perceptions for implementation and its effect on improving teacher practice.

**Personal Reflection**

I have served as a building administrator at the middle school level for the past fourteen years, four as an assistant principal and ten as a principal. I have also served on the executive boards of both the state and national Principals associations, as well as President of the Massachusetts Secondary Principal Association. In 2010, I was honored to be named the Massachusetts Middle School Principal of the Year, an award I share with the incredible staff at Raynham Middle School. I have also served on a committee for the past three years, with the Massachusetts Department of Elementary and Secondary Education, offering insight into the implementation and design of the evaluation system through each stage. In each of these key leadership roles at both the state and national level, I have endeavored, at every opportunity, to share my knowledge as well as learn from others. As such, my interest and research in other state evaluation models has lead me to believe the Massachusetts model, in its present application, is comprehensive, accurate, and fair, as well as very different from other states across the nation. In the near future Massachusetts will include student achievement measures as well as survey data from students and families in educator evaluations. The design and implementation of these aspects will further test the fairness of the system and ultimately its success.

Personally, I have seen the new evaluation system change the conversations in schools around teaching and learning, in turn increasing the focus on professional practice and student achievement. This process works best when there is respect and trust among the evaluators and those being evaluated. The focus of these partnerships needs to be focused in developing
teaching professionals, supporting and enabling their learning, so that these high standards of professional practice, as outlined in the Massachusetts teacher evaluation system, can be met. Considering the results of the survey data, one could conjecture that these conditions may have existed in District A. My hope is that conversations about high quality, effective teaching, as well as student achievement is supported through the new evaluation process. This focus on student outcomes and professional growth should enable educators, to reach more students and affect more positive student achievement in the future. As evaluators and educational leaders, we must remember the key components of performance improvement theory, namely the clarity of the goals and expectations as well as provide quality feedback to facilitate professional improvement. I do believe that this evaluation model can impact education positively; however, I also believe that its success rests on the ability of evaluators to foster trust, collaboration, and continued growth. In closing, my interest in this topic will not cease with this thesis. I intend to continue to lending my voice to a system of evaluation that can benefit educators and the students they serve.
References


Appendix A
Letter Requesting Permission to the Superintendent of Schools

January 3, 2013

Dear Superintendent,

I am currently enrolled in a doctoral program at Northeastern University and am in the process of completing the dissertation stage of the program. My research is focused on the new teacher evaluation system specifically teacher perspectives regarding its value and implementation. The study is designed to use a quantitative approach consisting of collecting survey data and focus groups from individuals who helped pilot the evaluation system in three separate districts. Therefore, I am requesting permission to elicit the participation of these recent participants in the early adoption project in your district.

The study will be conducted as an evaluation of the process and its impact on teacher practice. This information could possibly help in the implementation process moving forward and increase the effectiveness of the evaluation system to increase teacher effectiveness and student achievement.

My intention is for this evaluation study to benefit the implementation of the new evaluation system and help the administration and faculty maximize the possible positive effects of the program on the teaching profession and student achievement.

If you have questions regarding this study, please contact me directly at (774) 254-4442 or email me at Thomson.d@husky.neu.edu. You may also contact the chairperson of my committee, Dr. Christopher Unger at Northeastern University, (617) 909-1360. Thank you for your consideration and time. I look forward to talking with you in the future about my research.

Sincerely,

David L. Thomson
Doctoral Candidate, College of Professional Studies
Northeastern University, Boston
Principal, Raynham Middle School
Raynham, MA 02767
Appendix B
Initial Email to Participants

January XX, 2013

Dear Educators,

My name is David Thomson, and I am a doctoral candidate in the College of Professional Studies at Northeastern University and the Principal of Raynham Middle School in Raynham, Massachusetts. As part of my dissertation research, I am conducting a study on teacher perceptions of the new evaluation system, specifically its implementation, and its effect on instructional practice.

In order to gather data about this important and timely topic, I am inviting you to participate in my study by filling out a quick online survey. You have been asked to participate in this survey because you were part of the pilot of the evaluation system in your district. Due to your previous experience, you have a unique perspective on the process. Your responses to the survey will help me gain insight into teachers’ perceptions of the new evaluation system and its impact on instructional practice in the classroom.

The survey will take approximately 15-20 minutes to complete. You will receive an email invitation to participate in the survey, you will be presented with a quick overview and description of the purpose of the survey at the bottom of this email will be a link to the survey. By clicking on the link to the survey you will acknowledge your understanding of this information, consent to take the survey, and proceed to the actual survey questions. The survey will begin by asking what level you teach and the range of experience you have in the profession. Then, you will respond to survey questions by choosing a number to indicate the extent of your feelings about certain statements.

Your participation in this survey is completely voluntary. There is no punitive action if you choose not to participate.

The survey is completely confidential. If you have any questions, please contact David Thomson at (774) 254-4442 or email at thomson.d@husky.neu.edu. Thank you for your consideration.

Thank you,

David L. Thomson
Doctoral Candidate, College of Professional Studies
Northeastern University, Boston
Principal, Raynham Middle School
Raynham, MA 02767
Appendix C
Unsigned Consent Document for Online Survey

Northeastern University, College of Professional Studies

Investigator(s): Principal Investigator, Dr. Christopher Unger and Student Researcher, David Thomson

Title of Project: Teachers’ perceptions of the new Massachusetts teacher evaluation instrument and process on instructional practice.

Request to Participate in Research:
I would like to invite you to participate in a web-based online survey. The survey is part of a research study whose purpose is to examine the effect of the new teacher evaluation system on teacher perceptions and instructional practice. This survey should take about 15-20 minutes to complete.

I am asking you to participate in this study because you were evaluated as part of the early adopter, or pilot of the new evaluation system and possess prior experience with the system. You must be 18 years old to take this survey.

The decision to participate in this research project is voluntary. You do not have to participate and you can refuse to answer any question. Even if you begin the online survey, you can stop at any time.

There are no foreseeable risks or discomforts for you for taking part in this study, and your participation will not affect your teaching position in any way.

There are no direct benefits to you from participating in this study. However, your responses may help us learn more about the impact and implementation of the new teacher evaluation system. You will not be paid for your participation in this study.

Your participation in this study is confidential. There are a few questions that will ask for identifying information such as the level at which you teach and your length of time in the profession. However, given the number of expected participants, it will be very unlikely that any one participant could be directly identified. Because of the nature of web-based surveys, it is possible that respondents could be identified by the IP address or other electronic record associated with the response. Neither the researcher nor anyone involved with this survey will be capturing this type of data. Any reports or publications based on this research will use only group data and will not identify any individual or school as being affiliated with this project.

If you have any questions regarding electronic privacy, please feel free to contact Mark Nardone, IT Security Analyst via phone at 617-373-7901, or via email at privacy@neu.edu.

If you have any questions about this study, please feel free to contact David Thomson, the person mainly responsible for the research, by phone at 774-254-4442 or by email at Thomson.d@husky.neu.edu.
You can also contact Dr. Christopher Unger by phone at 617-373-2400 or by email at c.unger@neu.edu.

**If you have any questions regarding your rights as a research participant**, please contact Nan C. Regina, Director, Human Subject Research Protection, 960 Renaissance Park, Northeastern University, Boston, MA 02115. Tel: 617.373.7570, Email: irb@neu.edu. You may call anonymously if you wish.

Please complete this survey by *(insert appropriate date)*
**By clicking on the survey link below you are indicating that you consent to participate in this study. Please print out a copy of this consent form for your records.**

http://www.surveymonkey.com/s/DR3
Thank you for your time.
David L. Thomson
Appendix D
Survey

Online Survey

Years in Education

1-5  6-10  11-15  over 15

What grade level do you currently teach?

Elementary    Middle    High School

Please complete the following with the response that best represents your opinion of the statement regarding the new teacher evaluation process.

1. The self-reflection process, as prescribed in the new teacher evaluation model, caused me to seriously consider my teaching practice.

1-Strongly disagree  2-Disagree  3- Neither  4- Agree  5- Strongly Agree

2. The rubric, the standards and indicators of effective teaching practice, is a comprehensive guide of the traits of effective teaching.

1-Strongly disagree  2-Disagree  3- Neither  4- Agree  5- Strongly Agree

3. Setting SMART goals has helped me to focus my efforts on increasing student achievement.

1-Strongly disagree  2-Disagree  3- Neither  4- Agree  5- Strongly Agree

4. Setting SMART goals has helped me focus on improving my teaching practice.

1-Strongly disagree  2-Disagree  3- Neither  4- Agree  5- Strongly Agree

5. The new evaluation process resulted in changes in my teaching methods.

1-Strongly disagree  2-Disagree  3- Neither  4- Agree  5- Strongly Agree

6. My evaluator was able to gather sufficient evidence to accurately assess my performance.

1-Strongly disagree  2-Disagree  3- Neither  4- Agree  5- Strongly Agree

7. Looking back, the evaluation process resulted in positive changes in my professional practice.
1-Strongly disagree   2-Disagree   3- Neither   4- Agree   5- Strongly Agree

8. Would you be willing to take part in a follow up focus group discussion? IF so please share an email address where you can be reached.

Open Response

9. Overall what are your thoughts on the new teacher evaluation system?

10. What do you like about the new teacher evaluation system?

11. What do you not like about it?

Thank you for completing this survey!

As a follow-up to this survey, I will be conducting focus groups to learn more about the topic of my research. If you decide to participate, you will be asked to partake in a group discussion about your experiences with the new teacher evaluation system and its influence on your teaching practice. The session will take place in your district in one of the school buildings after school hours, and should last about one hour.

During the focus group session, you will be talking with 5-7 other teachers who have experienced the new teacher evaluation system. This discussion will be audio taped so that I can accurately reflect on what is discussed. I will be the only person reviewing the recordings and transcribing the discussion. No one else will have access to this data. After the study is complete the audio recordings and transcriptions will be destroyed. Please also be assured, that if you participate, you do not have to answer any questions that you don’t wish to.

Your participation as well as your responses will be held confidential. The information gathered from the group discussion will be kept in a secure location away from your district and school, furthermore your identity will not be revealed to others. However, the nature of a focus group is a group discussion, and it is possible that other group members could tell someone else. Because of this dynamic, I cannot promise that what you say will remain completely private, but I will ask that you and all other group members respect the privacy of everyone in the group.

Participation in this study is entirely up to you and one does not have to participate. You may also leave the study at any time. Although you probably will not benefit directly from participation in this study, this research is intended to improve the implementation of the teacher evaluation system. The results could potentially help improve the implementation of the evaluation system and may benefit teachers and students today and in the future.

Thank you for your consideration. If you would like to participate or have questions about this study, please email me at thomson.d@husky.neu.edu by January XX, 2013.
Thank you,

David L. Thomson
Doctoral Candidate, College of Professional Studies
Northeastern University, Boston
Principal, Raynham Middle School
Raynham, MA 02767
Appendix E
Survey Participation Email Reminder

January XX, 2013

Dear Participant,

My name is Dave Thomson, and I am a doctoral candidate in the College of Professional Studies at Northeastern University and the Principal of Raynham Middle School in Raynham, Massachusetts. I am conducting research about the new teacher evaluation system and its implementation and its effect on teaching practice.

You received an email recently that solicited your participation and contained a link to an online survey. For those of you who have not yet responded to the survey, I am sending this as a reminder. If you already filled out the survey, thank you so much for doing so! Your response to these questions will help me to gather important data that will enlighten others on the implantation and impact of the new teacher evaluation system, and may lead to improvements to the program.

The survey will take approximately 15-20 minutes. This survey is voluntary. If you would like to participate, click on the link below. The survey must be completed by January XX, 2013.

http://www.surveymonkey.com/s/DR3

Thank you for your time and consideration.

Sincerely,

David L. Thomson
Doctoral Candidate, College of Professional Studies
Northeastern University, Boston
Principal, Raynham Middle School
Raynham, MA 02767
Appendix F
Signed Informed Consent for Focus Group Participants

Northeastern University, College of Professional Studies

Investigator Name: David Thomson

Title of Project: The impact of the new teacher evaluation instrument on teacher perceptions of the evaluation process and instructional practice

Informed Consent to Participate in a Research Study
I am inviting you to take part in a research study. This form will tell you about the study, but the researcher will explain it to you first. You may ask this person any questions you have. When you are ready to make a decision, you may tell the researcher if you want to participate or not. You are not required to participate if you do not want to. If you decide to participate, the researcher will ask you to sign this statement and will give you a copy for your records.

Why am I being asked to take part in this research study?
You have been asked to participate in this study because you are a practicing teacher in an early adopter district that piloted the Massachusetts Educator Evaluation system last year. As an early adopter, you have increased experience with the new evaluation system and thus increased perspective.

Why is this research being done?
The purpose of this research is to explore the impact of the new evaluation system on teacher instructional practice and the process of evaluation.

What will I be asked to do?
The researcher will be looking for you to participate by taking part in a focus group that will be audio recorded.

Where will this take place and how much time will it take?
The focus groups will take place in a school building in your district afterschool; we anticipate the focus group to take approximately one hour.

Will there be any risk or discomfort to me?
There are no foreseeable risks involved in taking part in this study. All responses will be kept confidential, including participants’ names, schools, and districts. All participants in the focus group will agree to keep all information regarding participants and their responses confidential.

Will I benefit from this research?
There will be no direct benefit to you for taking part in the study. However, the information learned from this study could impact the future implementation of educator evaluation in the state of Massachusetts and beyond.
Who will see the information about me?
Only the researcher of this study will see the information about you. No reports or publications will use information that can identify you in any way. All audio recordings will be erased after analysis.

If I do not want to take part in this study, what choices do I have?
You are not required to participate in this study. Stopping your participation will not affect your professional standing. At any time during the study, you may refuse to answer questions, as well as end your participation. If you choose not to participate, do not sign and ignore this form.

Whom can I contact if I have questions or problems?
David Thomson
Principal Investigator
Cell #: (774) 254-4442
Email: thomson.d@husky.neu.edu

Chris Unger, Ed.D.
Overseeing Study
Campus #: (617) 337-2400
Email: c.unger@neu.edu

Northeastern University, Boston

Whom can I contact about my rights as a participant?
If you have any questions about your rights as a participant, you may contact Nan C. Regina, Director, Human Subject Protection, 960 Renaissance Park, Northeastern University, Boston, MA 02115 tel. 617-373-7570, email: irb@neu.edu. You may call anonymously if you wish.

Will I be paid for my participation?
There is no compensation for participation in this study.

Will it cost me anything to participate?
There is no cost to participate in this study.

I have read, understood and had the opportunity to ask questions regarding this consent form. I fully understand the nature and character of my involvement in this research program as a participant and the potential risks. Should I be selected, I agree to participate in this study on a voluntary basis.

___________________________________
Research Participant (Printed Name)

___________________________________
Research Participant (Signature) ____________ Date
Appendix G
Focus Group Questions

The following questions are to be asked of focus groups.

Opening Question:

Our purpose today is to discuss your perceptions of the new teacher evaluation system. We will be specifically discussing how this program was implemented and its impact on your instructional practice. In order to ensure confidentiality I will ask that everyone agree to not share any individual’s names, responses, or any other personal identifiers.

1. Looking back on the new evaluation process, what do you think was the most significant aspect of the new system and why?

2. How did this aspect effect you?

3. Tell me your impressions of the self-reflection process.

4. Tell me your impressions of the goal-setting meeting with your evaluator.

5. What are your opinions of the rubric; the standards and indicators of effective teaching practice?

6. Were you familiar with all of the terms and indicators found in the rubric?

7. Do you feel you had enough background information, or were provided with the information to achieve proficiency in all categories?

8. Describe your mid cycle and summative reviews with your evaluator. Did you find the feedback helpful in reaching your goals and improving your practice?

9. Do you feel enough evidence of practice was collected to accurately assess your teaching performance?

10. Were there any changes in your teaching practice that you attribute to the new system? Can you give an example?

11. Overall, do you feel the evaluation system has the capacity to improve student achievement? How?
Appendix H
Focus Group Participation Email Reminder

January XX, 2013

Dear Survey Participant,

I would like to take this opportunity to thank you for completing the survey that contributes to my research on teacher perceptions of the new Massachusetts teacher evaluation system. Your participation has supplied data that will help to evaluate these initial steps of implementation of this groundbreaking evaluation system.

I would like to ensure you are aware of an opportunity to participate in focus group discussions that will contribute further to this research. I would like to take this opportunity to elicit and encourage your participation. If you are interested in participating in a focus group, please email me within the next three days at thomson.d@husky.neu.edu. If you are able to participate, you will be asked to partake in a group discussion with approximately 7-9 other teaching colleagues who also piloted the evaluation system in your district. In particular, you will be asked about how this program was implemented and how it affected your instructional practice. Although we will be meeting in a group, we will discuss and agree to hold individual names, accounts, and identifiers confidential. As the sole recorder of these discussions, all transcriptions will be void of names and other personal identifiers. Once the discussion is transcribed, the audio recordings will be destroyed. These sessions will take place after school, in a school building in your district and should last approximately one hour.

As I have mentioned before, participation in this study is entirely up to you. However, I would ask that you consider the contributions this research could add to the future implementation of the evaluation system. If you do decide to participate, please contact me at thomson.d@husky.neu.edu.

Thank you,

David L. Thomson
Doctoral Candidate, College of Professional Studies
Northeastern University, Boston
Principal, Raynham Middle School
Raynham, MA 02767
Certificate of Completion

The National Institutes of Health (NIH) Office of Extramural Research certifies that David Thomson successfully completed the NIH Web-based training course "Protecting Human Research Participants".

Date of completion: 12/02/2011

Certification Number: 810902