BRIDGEWATER STATE UNIVERSITY

The Graduate Review

Volume 3 Article 15

2018

The Massachusetts Educator Evaluation System and Teacher Perceptions of Professional Growth

Craig Goldberg
Bridgewater State University

The Massachusetts Educator Evaluation System and Teacher Perceptions of Professional Growth

CRAIG GOLBERG

Introduction

In 2011, the Commonwealth of Massachusetts adopted new guidelines for the evaluation of all teachers. These regulations, which are state law, incorporate traditional elements such as administrator evaluation of educators through observations and student learning data. The system also takes into account modern trends of educator evaluation such as teacher-led analysis, reflection, planning, action steps, collaboration, and the use of a standards-based rubric.

This system places on teachers ownership of providing evidence of proficiency as they are required to display aptitude related to several major standards and indicators such as curriculum, planning, assessment, and professional practice. Another element of this system is showcased when teachers create student learning and professional practice goals and offer evidence of their proficiency in meeting them.

While this new system considers traditional aspects as well as modern trends in educator evaluation, it is not without criticism. Some educators and teacher unions are skeptical of use of student data related to state-wide assessment acts, in part, as an evaluative tool. However, not all educators teach subjects that have state assessments. Regardless, assessment data as an evaluative tool will not take into account the group of students one teaches (i.e., students with disabilities, English language learners, and other learning challenges) or the socioeconomic status of the district, and thereby the caliber and the expectations asked of the students.

Another potential issue is the amount of time, effort, energy, and money needed to implement the procedures and processes. From a teacher's perspective, creating evaluation evidence, which prove they are meeting the state's standards, is extremely time-consuming.

A final potential problem with the new evaluation system mandated by the Commonwealth of Massachusetts involves the ability to accomplish its goals. Teachers and administrators feel the burden of all that is asked of them from the federal government, the state, and their individual districts. They question whether adding another item to their list will turn teachers away from the profession all together, potentially causing talented individuals to go into other fields.

Purpose Statement

The purpose of this study was to examine the effect of State-mandated teacher evaluation on teacher perceptions of professional growth.

Research Questions

1. To what extent do teachers feel the teacher evaluation rubric is reflective of quality teaching?
2. What is the relationship between the state-mandated evaluation process and teacher perceptions of their professional growth regarding curriculum?
3. What is the relationship between the state-mandated evaluation process and teacher perceptions of their professional growth regarding planning?
4. What is the relationship between the state-mandated evaluation process and teacher perceptions of their professional growth regarding assessment?
5. What is the relationship between the state-mandated evaluation process and teacher perceptions of their professional growth regarding goals?

Literature Review

The literature review reveals that increased federal involvement in education policy over the past sixty years have mandated a fusion of high-stakes tests and educator evaluation. Groen (2012) argues this was done for both educational and political reasons. For example, as an incentive to comply with desegregation orders, the federal government provided local school districts money to fulfill their legal obligations, while also promoting educational programs to assist underprivileged students. With the No Child Left Behind Act being effective in 2001, high-stakes testing and teacher accountability were incentivized across the nation.

Teacher evaluation continues to be a major tenet of the educational-reform movement. The federal Race to the Top Initiative has spurred development and implementation of new teacher evaluation systems as a key lever for improving school effectiveness and raising student achievement (Master, 2013). Described by Sabol (2013) as a "seismic shift in the educational landscape" (p.13), student assessment results have become a central indicator of learning, and their inclusion in many teacher evaluation tools is common.

Doherty and Jacob (2013) noted that 48 states have implemented varying levels of evaluation systems, with 45 of those states requiring formal observations as an evaluative tool. They reveal two main approaches to teacher evaluation reform: value-added measures and standards-based evaluations. Value-added models attempt to estimate a teacher's contribution to student test-score growth. In contrast, standards-based evaluations take into account rigorous and data-driven classroom observations in which evaluators assess a teacher's practice relative to explicit and well-defined district standards.

Papay (2014) argues that the value-added model has many limitations, causing a negative effect on teacher evaluation. With standardized testing in place in most states, fewer than one in three teachers work in a grade or subject area that supports value-added analysis as state assessments typically only include English language arts (ELA) and mathematics. While the val-

ue-added model offers hard data, standards-based evaluations can be more objective, as the evaluator bases conclusions in a qualitative fashion. Standards range from student achievement to professional responsibilities.

One study found that standards-based evaluations helped teachers become more reflective and focused on their teaching (Montecinos et al., 2010). They found that the use of standards-based performance indicators and rubrics made "the process transparent and allow faculty to develop a common understanding of what quality teaching means" (p. 287). If a teacher did not meet a standard, the rubric clearly showed steps towards improvement in a self-directed fashion to meet the standard in the next evaluation cycle. To further support standards-based evaluations, Master (2014) found in a study that teachers valued an administrator's holistic judgment because these statements can capture aspects of job performance, such as teaching to diverse learners, that may be missed by more evaluation instruments.

Teachers need to feel motivated if they are going to "buy-in" to teacher evaluation. Firestone (2014) argues that if teacher evaluation is going to be effective, teachers need to feel motivation based on intrinsic factors rather than external motivation (such as based on pay and prestige). Conducting research on motivation theory related specifically to teachers, Firestone determined several key school-based elements necessary to intrinsically motivate teachers to see evaluation as a tool in their professional growth, such as providing teachers with useful feedback.

The new mandated teacher evaluation tool in Massachusetts is teacher-centered as it relies on educators to create a portfolio of work showing proficiency in a set of standards as well as individual goals. However, Thompson (2014) studied teachers' perceived professional growth as a result of Massachusetts' new teacher evaluation system in three early adopter districts. He found that the new evaluation process had veteran teachers perceiving less professional growth than novice teachers. He concluded that the results of the study in early adopter schools showed that the impact of educator evaluation as perceived by teachers was rather mixed.

Methodology

This was a quantitative, non-experimental, descriptive study, using a descriptive survey as an instrument to attain data. This study was intended to examine the effect of state-mandated educator evaluation on teacher perceptions of professional growth.

Setting

The setting for this study included six high schools in suburbs south of Boston, Massachusetts. The schools shared similar homogeneous demographics and community socio-economic status. Each school was located in a community that was predominately middle to upper-middle class and largely Caucasian. The median income of the six school districts was $97,000, and the average ethnic make-up was 93% Caucasian.

The six schools totaled a possible respondent pool of

723 teachers. School A had a total teacher population of 109. School B had a teacher population of 132. School C had a total teacher population of 132. School D had a total teacher population of 60. School E had a total teacher population of 153. School F had a total teacher population of 136. The researcher anticipated a response rate of 35%.

Participants

Respondents included teachers holding a valid teaching license in the Commonwealth of Massachusetts, and who worked currently under each district's teacher collective bargaining agreement. This included educators who taught core subjects such as mathematics, ELA, history, science, and foreign languages. This also included special education teachers, physical education teachers, and those who taught specialties such as business, art, and industrial technology courses. The participants had a varied amount of teaching experience.

Instrumentation

This survey was created by the researcher (see Appendix) and distributed using Google Forms, incorporating elements of the Massachusetts Educator Evaluation system and the research questions. The variables, which addressed the research questions, included elements of quality teaching, curriculum, planning, and assessment. Further variables were created, incorporating elements of goal setting and usefulness of the state's teacher rubric. The survey instrument was reviewed by each school administrator prior to agreement to participate in this study.

Page one of the Google Forms document sent to possible respondents contained an explanation of the study and an introduction thanking participants for taking the time to respond to the survey. Page two of the survey contained questions regarding demographic information such as school, years of teaching service, and current evaluation rating. Page three contained questions that aligned with the research questions. Respondents were instructed to use a Likert Scale of strongly disagree (1), disagree (2), agree (3), and strongly agree (4) to respond to these questions. An optional comment box was also available for respondents to leave feedback and general comments.

Procedure

The first step taken to implement the study was to recruit principals from schools in the Boston suburbs who shared similar homogenous demographics. The researcher sent e-mails, called schools, mailed information packets, and utilized contacts from a list of 27 prospective schools, of which 6 principals agreed to allow their staff to participate in the study.

Google Forms, a web-based survey software, was used as a method to distribute the survey in mid-September 2016 as the principals of each participating high school were sent an e-mail that contained information about the study, the researcher, and a link to the Google Forms online survey. This e-mail also contained information ensuring that teachers were aware that the survey was anonymous and completely optional. The survey was then forwarded to each respective school's staff. Respondents had two weeks to complete

the online survey before submissions were no longer accepted.

The researcher used the Statistical Package for the Social Sciences (SPSS) for statistical analysis of data. The data from the completed surveys were coded and entered into SPSS in October 2016. The researcher analyzed Mean data, standard deviation, and independent samples t-Tests for several months to determine statistically significant findings.

Findings

There were 173 total respondents in this study, with a response rate of 24%. Demographically the following profile emerged from six participating high schools. Thirty-one percent of the respondents taught at School A, fifteen percent at School B, eleven percent at School C, eight percent at School D, twenty-seven at School E, and thirty-three at School F. Eleven percent of the teachers taught mathematics, fifteen percent science, eighteen percent social studies, sixteen percent ELA, eight percent a foreign language, thirteen percent in special sducation, and the remaining twenty percent responded to the option "other". This category included business, media production, art, physical education, and industrial arts. Sixty-six percent of the respondents were female, and thirty-four percent were male.

Educators are evaluated by school administration. Twenty-three percent were evaluated by their principal, forty-two percent by an assistant principal, thirty percent by their department head, and the remaining five percent responded with the "other" option. Those answering "other" noted "district superintendent", "it changes annually", and "any of the above".

The respondents were well educated and had a great deal of experience in the field of education. Thirteen percent earned only a Bachelor's Degree, sixty-four percent earned one Master's Degree, eleven percent earned more than one Master's Degree, nine percent earned a Certificate of Advanced Graduate Studies (CAGS), and three percent earned a Ph.D. Twelve percent of respondents had been teaching between one-five years, twenty-one percent had been teaching between six-ten years, twenty-two percent had been teaching between eleven-fifteen years, and the remaining forty-five percent had more than sixteen years teaching experience. Sixty-eight percent stated they took on professional responsibilities outside of teaching such as coaching a sport and/or advising a club, while thirty-four percent did not.

Further demographics data included 86% of teachers having earned professional status in their district, while 14% did not. Underscoring their high level of experience, 82% of the teachers were on a two-year, self-directed growth plan that is reserved for those with both professional status and three or more years teaching experience in a particular school. Eighteen percent were on a one-year, self-directed growth plan that is reserved for newer teachers. Using the state and/or their district evaluation ratings, 84% earned a rating of proficient, and 16% were rated as exemplary. None of the teachers currently held a rating of unsatisfactory or needs improvement. Regarding evaluation ratings, 68%

of the teachers felt that evaluators should have the option to rate them between these rating levels, while 32% felt the current rating system was fair as it is.

Teachers were asked if they had experience with other evaluation systems in the past, and if the new evaluation system offered a marked improvement from previous systems of educator evaluation. Only 13% of respondents had experience only with this system, indicating they are new to the field of education. As 87% of respondents had 6 or more years of experience, they had undoubtedly had experience with other evaluation systems in the past. However, only 17% of respondents stated this new system was an improvement from previous systems, while 47% stated it was not better, and 23% were not sure.

By and large, teachers did not perceive that they were gaining professional growth as a result of the new evaluation system. One of the more straightforward items in the survey inquired about teacher perceptions of receiving professional growth as a result of the new educator evaluation system. The scale of responses was: strongly disagree (1), disagree (2), agree (3) and strongly agree (4). Teachers did not feel that the educator evaluation system helped them in their professional growth (n=173, M=2.11, SD=.75).

Furthermore, Mean data illustrate that respondents were neutral to the statement that the evaluation system caused change to their professional practice (n=173, M=2.54, SD=.75). For many variables, teacher responses about the growth they experienced as a result of the evaluation rubric produced a Mean of less than 2.50, except where otherwise distinguished.

There were no noteworthy statistically significant differences in any of the dependent variables when assessed by the independent variables of gender, school size, professional status, and whether a teacher advised an after-school club or coached a sport. Analyses conducted for the six individual schools and by department-subject produced mixed results.

The first research question examined "the extent that the teacher evaluation rubric was reflective of quality teaching". In responding to the statement that "the teacher rubric is a comprehensive guide of the traits of effective teaching", there was more dissatisfaction than satisfaction with the teacher rubric (n=173, M=2.38, SD=.82). However, respondents in schools that used the state's model rubric perceived it more poorly (n=70, M=2.31, SD=.88) than those in schools that had created their own adapted rubric (n=50, M=2.50, SD=.76).

The second research question focused on "the relationship between the state-mandated evaluation process and teacher perceptions of professional growth regarding curriculum". Respondents tended toward disagreement that educator evaluation caused changes to the delivery of their curriculum (n=173, M=2.31, SD=.76).

The third research question investigated "the relationship between the state-mandated evaluation process and teacher perceptions of professional growth

regarding planning". The full sample of respondents (n=173), in treating these dependent variables individually, were toward disagreement: "I find myself collaborating with colleagues more since the implementation of the new evaluation system" (M=2.19, SD=.73); "Since implementing the new evaluation system, I have created more rigorous standards-based units" (M=2.30, SD=.68); "I find myself making more creative lessons since the new evaluation system was implemented" (M=2.01, SD=.69); and "The new evaluation system has caused me to think deeper about my lesson planning" (M=2.05, SD=.80).

The fourth research question evaluated "the relationship between the state-mandated evaluation process and teacher perceptions of professional growth regarding assessment". Respondents were asked if the new evaluation system had caused them to evaluate student data more than they had in the past, to which they tended toward disagreement (n=173, M=2.30, SD=.79). Furthermore, respondents disagreed that they were creating more non-traditional assessments to earn a proficient evaluation rating (n=173, M=2.15, SD=.73). Finally, teachers did not feel the evaluation system caused them to rethink the way they assess students (n=173, M=2.13, SD=.76).

The fifth research question concerned "the relationship between the state-mandated evaluation process and teacher perceptions of professional growth regarding setting goals". The Mean for these two dependent variables was contradictory: while teachers did give thought to their goals (n=173, M=2.77, SD=.78), they did not feel that this step helped them to focus on improving their practice (n=172, M=2.32, SD=.75).

Embedded in the survey instrument were additional questions about respondents' general experience with the evaluation system not connected to the primary research questions of the study but still worthy of consideration. These variables contributed to understanding the general effect of the Educator Evaluation System on teacher perceptions of professional growth.

Self-reflection is a critically assumed goal of the Educator Evaluation System in Massachusetts. As such, a successful evaluation system is expected to cause educators to be more reflective about their teaching practice. However, when asked if they have become more reflective as a result of the evaluation system and process, teachers responded toward disagreement (n=173, M=2.28, SD=.76). Furthermore, respondents were barely neutral that reflection led them to consider their teaching practice (n=173, M=2.40, SD=.75).

While there appears to be a general dissatisfaction with this evaluation system, total years of experience in education generated many significant differences among teachers in their response to the evaluation system. Those with 1-5 years of teaching experience (n=21) appeared to gain professional growth from the system. For example, regarding the effectiveness of the teacher rubric, a statistical difference resulted in the comparison of new teachers, those who have between 1-5 years experience (n=21, M=2.76, SD=.77) and those with slightly more experience, 6-10 years (n=36,

M=2.25, SD=.87). A t-Test comparing these groups on this variable produced a statistically significant difference (t=.09, p<.03).

The only demographic group that felt the goal-setting process tended to help improvement in teaching were those with 1-5 years experience (n=27, M=2.71, SD=.78). A comparison with those who have 6-10 years of experience (n=40, M=2.30, SD=.79) on this variable resulted in a moderate difference (t=1.89, p<.05). Those with 1-5 years were then compared to those who had 11-15 years experience (n=32, M=2.32, SD=.66). Between these two groups there was also a moderate difference (t=2.07, p<.04). Finally, when the least experienced teachers were compared on this variable of whether the evaluation system produced improvement in teaching with those who had 16 or more years experience (n=79, M=2.23, SD=.75), the results indicated a strong statistical difference (t=2.58, p<.01).

The researcher transformed the variables of years teaching into two new variables: 1-5 years teaching experience and 6 and more years for the variable "I feel I have received professional growth as a result of the new Massachusetts Educator Evaluation system". Respondents who had been teaching 1-5 years were neutral about their professional growth due to the educator evaluation system (n=21, M=2.57, SD=.68). This was compared to the disagreement of those with more than 6 years experience (n=152, M=2.05, SD=.74) about whether the evaluation system produced professional growth. A t-Test comparing the Mean of the newer teachers with this combined group of more experienced teachers about perceptions of professional growth produced a strong statistically significant difference (t=.82, p<.00).

Respondents with 1-5 years of experience found reflection to occur as a result of the evaluation process (n=21, M=2.76, SD=.70), while those with 6 or more years of experience did not (n=152, M=2.21, SD=.75). Comparing years of teaching experience produced strong statistically significant results on this variable (t=3.12, p<.00).

Teachers with only a Bachelor's Degree (n=22, M=2.59, SD=.80) were slightly greater than neutral when asked if the evaluation system assisted in changes to their delivery of curriculum. Those with a Master's Degree were close to disagreement responding to this variable (n=112, M=2.22, SD=.76). A t-Test comparing teachers with these two levels of education and their perceptions that the evaluation system effected changes in their delivery of the curriculum resulted in a statistically significant difference (t=.77, p<.04).

Further analysis of teachers with less experience utilized a comparison by educational level in the perception of growth as a result of educator evaluation. Those with only a Bachelor's Degree (n=22, M=2.63, SD=.79) were compared to the transformed demographic those with a Master's Degree or higher (n=151, M=2.32, SD=.74) on this variable of whether they have become more reflective as a result of the evaluation system. A t-Test comparing these Means resulted in a statistically significant difference (t=2.36, p<.01).

Comparisons of teachers with experience with other evaluation systems against teachers with experience only with the new system produced some significant findings. Respondents with only experience with the new system were generally neutral that it impacted their professional growth (n=23, M=2.43, SD=.73). However, those who had experience with other systems reported strong disagreement that the new system is better than previous evaluation methods (n=150, M=2.00, SD=.75), producing statistically significant results (t=2.20, p<.00).

Regarding the degree to which the self-reflection process as connected to evaluations was causing teachers to consider their teaching practice, it appears that those who had only experience with the new system were generally neutral to the statement (n=23, M=2.60, SD=.84). However, individuals with experience with other systems, and did not feel the new one was better, reported that the evaluation system was not causing them to make changes in their teaching practice (n=150, M=2.09, SD=.73), displaying statistically significant results (t=5.50, p<.00).

The role of the teacher's primary evaluator was also addressed. When considering the new style of mini-observations conducted by evaluators, respondents tended to agree that they preferred them more than traditional, full-period observations (n=173, M=2.68, SD=.84). Concerning the usefulness of formative meetings, respondents were neutral that these sessions were important to their professional growth (n=173, M=2.46, SD=.83). An identical result was found in the teacher evaluation of whether summative meetings were important to their professional growth (n=173, M=2.46, SD=.84).

Those teachers whose primary evaluator was a department head found the value in terms of professional growth very low, that is greater disagreement to the variable "I feel I have received professional growth as a result of the new Massachusetts Educator Evaluation system" (n=51, M=1.94, SD=.76). However, teachers appeared to perceive the educator evaluation process less poorly in terms of the professional growth experienced when their primary evaluator was the school's principal (n=40, M=2.37, SD=.74) than when the primary evaluator was an assistant principal (n=72, M=2.04, SD=.72). It is important to note that each demographic group had a mean less than 2.50, indicating a tendency toward, and, in two cases, outright disagreement that the evaluation system had caused them to perceive professional growth.

When teachers whose primary evaluator was a principal were compared to those who primary evaluator was an assistant principal, there was a statistically significant difference (t=2.32, p<.02). Furthermore, when teachers whose primary evaluator was a principal were compared to those whose primary evaluator was a department head an even greater statistically significant difference was found (t=2.73, p<.00).

All the respondents in this study were rated as either exemplary (n=28) or proficient (n=145) teachers. Principally, respondents rated as "exemplary" per-

ceived the evaluation system somewhat favorably compared to those who earned a "proficient" rating. Those who were rated "exemplary" viewed at a level of stronger agreement (n=28, M=2.57, SD=.69) than those who were rated "proficient" (n=145, M=2.25, SD=.80) regarding the variable of teachers using assessment data to drive instruction. A t-Test comparing groups on this variable resulted in a moderate statistically significant difference (t=2.38, p<.01).

A teacher's most recent evaluation rating also had a perceived impact on changes to teaching methods. Those who were rated "exemplary" (n=28, M=2.60, SD=.57) perceived more changes to their instructional methods than individuals who were rated "proficient" (n=145, M=2.20, SD=.72). A t-Test found very strong significant results (t=2.81, p<.00).

Additionally, those teachers rated as "exemplary" did not see educator evaluation as a necessary evil (n=28, M=2.35, SD=.83), while those whose evaluation rating was "proficient" did (n=145, M=2.73, SD=.75). A comparison of the two groups on this dependent variable showed a statistically significant difference (t=2.38, p<.01).

Conclusion

Most respondents in this study had many years of teaching experience, took on after-school activities, worked toward advanced degrees, and had professional status in their schools all to earn high teacher ratings. In other words, it is reasonable to state that this group of teachers were veteran, professional, and valued by their school community. Inferences and conclusions about the degree of the connection between the Massachusetts Educator Evaluation system and process, and how the teachers conducted their professional lives are provisional but also shed some light about how worthwhile this initiative is five years into its existence.

The data in this study strongly suggest that the Massachusetts Educator Evaluation system has not affected positive change in most educators' professional growth in any significant way. Teachers responded negatively when asked directly if they felt the new evaluation system caused them to experience professional growth. Respondents felt similarly when asked if their evaluation caused any changes to their professional practice. Consequently, the teacher response of this sample became a matter of how displeased teachers expressed themselves as being with the new system, than the degree of how much growth they were experiencing because of it.

There were areas in this study where significant differences were expected at the outset of the study, but do not appear to be confirmed by the evidence. First, there were no differences by gender; both male and female teachers held negative views of educator evaluations' impact on their professional growth. Also, though it was thought possibly otherwise, there was no difference comparing respondents who were involved in after-school activities, and those who were not. It might logically be presumed that educators who dedicated time to advise clubs or coach sports might be more invested in their professional practice. Howev-

er, both groups reported little or no impact because of state-mandated evaluation.

In analyzing the data pertaining to the first research question about the efficacy of the teacher rubric, there appears to be no significant connection between aspects of the teacher rubric such as curriculum, planning, assessment, and goals with increased professional growth in most educators. Generally, respondents felt more dissatisfaction than satisfaction with the rubric.

While the teacher rubric was created by the state, districts had the option to use the state's rubric as a model to fashion their own. Data in this study suggest that teachers who used the state's rubric had slightly more dissatisfaction than districts that created their own. This conceivably can be attributed to the state rubric not being particularly user-friendly and being cumbersome to read. Districts that utilized the option to create their own rubric considered teachers as stakeholders in the development process. Districts tended to know their teachers better than the state does and could tailor the rubric to address their needs.

Regarding the researcher's questions concerning educator evaluations' effect on planning, curriculum, and assessment, educators throughout this study appeared not to feel any more reflective in their practice as a result of the evaluation process, but this does not mean reflection is not occurring. Reflection is a hallmark of the educator evaluation system in Massachusetts and is a term that is used frequently throughout the teacher rubric. Thus, the evaluation system does not appear to be assisting in this aspect of professional growth in this sample of teachers.

Respondents in this study appeared more receptive to their evaluator being higher-level administrators, mainly principals, but even assistant principals, than department heads. It can be reasoned that this is due to the weight of a school leader having direct discussions, particularly with newer educators. Respondents were asked if they perceived professional growth due to educator evaluation, and those whose primary evaluator was the principal provided the most favorable results compared to those whose evaluators were assistant principals and department heads. The data clearly show that educators care about their evaluation, so perhaps being evaluated by a principal is a way for teachers to show their principal they are doing a "good job".

However, many variables related to the primary evaluator still showed a negative view of teacher evaluation and were neutral to the impact of formative and summative meetings on their professional growth. Juxtaposed to teachers caring about their evaluation, they appeared to resent having to prove that they are good teachers to their evaluators, particularly the group of those who have been teaching more than 16 years. It is reasonable to conclude that many experienced teachers felt they had the least to learn through evaluation because they were self-motivated and skilled enough to change their teaching practices if they feel it was needed.

As noted previously, the teachers in this study

were all rated "proficient" and "highly effective", yet there were significant differences when comparing teachers with these two evaluation ratings. The data show that respondents who earned an "exemplary rating" used assessment data to develop instruction and perceived changes to their teaching methods and did not see educator evaluation as a necessary evil. This perhaps indicates that those who earned an "exemplary" rating do in fact work the hardest to embrace the evaluation process.

The findings appear to indicate years of experience and highest degree earned were a major influence on perceptions of professional growth due to educator evaluation. The data reliably indicate that teachers with 1-5 years of experience and holding only a Bachelor's Degree (though the smallest group demographically, n=21) were the only group to report that they gained professional growth from the evaluation process. This group had generally favorable views of the teacher rubric and of the impact on their evaluation of curriculum, planning and assessment.

This perception likely resulted from the possibility that those with less experience in education benefitted from feedback and perhaps were more prone to learn from their mistakes. Several newer teachers commented that they benefitted from having conversations with their evaluators. As teachers created their own evidence to show they had met evaluation standards, this may have benefitted new teachers as it caused them to reflect on their own practice.

While newer teachers appeared to benefit to the largest degree compared to their peers from the evaluation system, most respondents, especially those with six or more years' experience, had strong disagreement that they gained growth from this process. This corroborates results found in a similar study by Thompson (2014) in Massachusetts Educator Evaluation system early-adopter districts. This group had experience with other systems in the past, and while the Massachusetts system reflected modern trends in educator evaluation, they did not feel it was an improvement over what they had previously experienced. Comments by respondents suggest beliefs that they already did their jobs the best they could, and that they resented a message sent from the state that they were "guilty until proven innocent", meaning they inferred that the evaluation process made the assumption that they were not doing their job, and they needed to prove that they were. Furthermore, those with experience with other evaluation systems had seen past methods of evaluation come and go. Therefore, it is reasonable to state that they felt this new system will not have any longevity.

It may benefit educational leaders to consider how teacher evaluation is packaged prior to it being implemented. While school leaders cannot change what is handed to them by the state, they can shape the way it is implemented in their school. The data indicate, and respondent comments confirm that this group of educators felt this system is cumbersome and time-consuming. Respondents noted that the time they spend being observed by and having discussions with their evaluator was not worth the time it took to put together evaluation materials. Educational leaders can benefit by offering

additional professional development on aspects of the Educator Evaluation system to show further investment in their staff's professional growth.

Furthermore, this study provided insights about who may gain the most from educator evaluation. Newer teachers perceived the most growth as they may be most receptive to the structured encouragement and engagement in reflection and discussion about teaching. For this group, it is thus important that their primary evaluators should be school principals.

Finally, this study indicated that veteran teachers, i.e., individuals with six or more years' experience, did not perceive that they experienced professional growth as a result of participation in the Educator Evaluation system. These individuals had advanced degrees, experience with other evaluation systems, and possessed years of experience to guide their teaching. The state should strongly consider decreasing the frequency, or possibly for the most experienced, waiving the evaluation process altogether for these veteran teachers. There may be other creative ways in which the experience and expertise of these more veteran teachers can serve as models for novice teachers to grow, rather than forcing veteran teachers to prove and reprove their value in a school building.

References

Firestone, W. (2014). Teacher evaluation policy and conflicting theories of motivation. *Educational Researcher*, 43(2), 100-107.

Doherty, K. M., & Sandi, J. (2013). Connecting the dots: Using evaluations of teacher effectiveness to inform policy and practice. Washington, D. C.: National Council on Teacher Quality.

Groen, M. (2012). NCLB--The educational accountability paradigm in historical perspective. *American Educational History Journal,* 39(1/2), 1-14.

Master, B. (2014). Staffing for success: Linking teacher evaluation and school personnel: Management in practice. *Educational Evaluation & Policy Analysis*, 36(2), 207-227.

Montecinos, C., Rittershaussen, S., Solis, C. M., Contreras, I., & Contreras, C. (2010). Standards-based performance assessment for the evaluation of student teachers: A consequential validity study. *Asia-Pacific Journal of Teacher Education*, *38*(4), 285- 300.

Papay, J. (2012). Refocusing the debate: Assessing the purposes and tools of teacher evaluation. *Harvard Educational Review*, *82*(1), 123-141.

Sabol, F. (2013). Seismic shifts in the education landscape: What do they mean for arts education and arts education policy? *Arts Education Policy Review*, *114*(1), 33-45.

Thompson, D. (2014). *Teachers' perceptions of the new Massachusetts teacher evaluation instrument and process on instructional practice* (Doctoral Disser-

tation). Retrieved from Northeastern University.

About the Author

Craig Goldberg began his teaching career for the New York City Department of Education after receiving a Bachelor's Degree from St. Joseph's College in Patchogue, NY and a Master's Degree at the State University of New York at Stony Brook. He graduated from Bridgewater State University in May 2017, having earned a C.A.G.S. in Educational Leadership. This research was completed under the mentorship of Dr. Stephen J. Nelson during the 2016-2017 academic year. He is currently a history teacher at Oliver Ames High School in Easton, Massachusetts. He lives in Easton, Massachusetts with his wife, Meredith; daughter, Madison; and son, Declan.

Appendix

Massachusetts Educator Evaluation System Survey

Instructions: Thank you for volunteering to respond to this 10-minute survey about the Massachusetts Educator Evaluation system. Although you may not personally benefit, this study is important because teacher feedback is essential to the success of any educator evaluation system. There are no foreseeable risks, your responses are anonymous (this form will NOT automatically collect your email address), and you may refuse to answer particular questions or withdraw from this survey at any time.

Please respond to one answer for each of the following questions regarding your experience with the Massachusetts Educator Evaluation system. If you have any questions, feel free to contact Craig Goldberg (C.A.G.S. in Education Leadership Student and Graduate Research Assistant) at cgoldberg@student.bridgew.edu. Thank you for your time in completing this survey. Please click "continue" to begin.

Part 1: Please answer the following questions. When you are done, click "continue" at the bottom of the page to go on to the next set of questions.

1. Which of the following schools do you currently teach at?
   - School A
   - School B
   - School C
   - School D
   - School E
   - School F

2. Of which department are you a member?
   - Math
   - Science
   - Social Studies
   - English
   - Special Education
   - Foreign Language
   - Other (please specify):

3. What is the student population of your building?
   - Under 1000-1250
   - 1251-1500
   - 1501-1700
   - 1701 or more

4. How many years have you been a teacher (please include experience at other schools, if applicable)?
   - 1-5
   - 6-10
   - 11-15
   - 16 or more

5. What is your highest degree earned?
   Bachelor's
   Master's
   More than 1 Master's
   CAGS
   Ph.D.

6. What is your gender?
   Male
   Female

7. Did school and/or district administration invest professional development time to explain to staff the Massachusetts Educator Evaluation system?
   Yes
   No
   Other:

8. Are you an adviser to an after-school club or coach a sport at your school?
   Yes
   No

9. For teachers who have been evaluated using different models in the past, do you feel the Massachusetts Educator Evaluation system is an improvement from previous methods of evaluation?
   Yes
   No
   Unsure
   I only have experience with the current Educator Evaluation system

10. Who is your primary evaluator?
   A Principal
   An Assistant Principal
   A Department Head
   Other:

11. The educator plan I am currently on is
   1 Year, Self-Directed Growth
   2 Year, Self-Directed Growth
   Directed Growth Plan (up to one year)
   Improvement Plan (30 days to one year)

12. Based on your most recent evaluation, what rating were you assigned by your evaluator?
    Exemplary
    Proficient
    Needs Improvement
    Unsatisfactory

13. Do you feel evaluators should have the ability to rate educators between rating levels (ex: between "proficient" and "exemplary")?
    Yes
    No

14. Have you earned Professional Status as an educator in your district?
    Yes
    No

15. Are you required to set a professional practice and student learning goal at the beginning of each evaluation cycle?
    Yes, I am required to set both goals
    I am required to set a Professional Practice goal only
    I am required to set a Student Learning goal only
    No, I am not required to set either

16. Does your district use a rubric to evaluate educators in your school?
    Yes
    No

17. Does your school use the rubric provided by the state or create its own rubric adapted from the state's?
    We use the state's model rubric
    We have our own rubric adapted from the state's model rubric
    Unsure

18. My district uses District Determined Measures (DDMs) or Common Assessments, and they have an impact on my evaluation.
    Yes
    No
    Unsure
    Other:

Part 2: In this next section, please rate whether you strongly disagree, disagree, agree, or strongly agree with each of the following statements. When you are finished, please click "continue" to go on to the final page.

19. Since the new educator evaluation system was implemented, I find myself evaluating student data to adjust my practice more than I did in the past.
    Strongly Disagree
    Disagree
    Agree
    Strongly Agree

20. The teacher rubric is a comprehensive guide of the traits of effective teaching.
    Strongly Disagree
    Disagree
    Agree
    Strongly Agree

21. The new evaluation process has resulted in changes to my teaching methods.
    Strongly Disagree
    Disagree
    Agree
    Strongly Agree

22. I feel I have received professional growth as a result of the new Massachusetts Educator Evaluation system.
    Strongly Disagree
    Disagree
    Agree
    Strongly Agree

23. My evaluator was able to gather sufficient evidence to accurately rate my effectiveness as a teacher.
    Strongly Disagree
    Disagree
    Agree
    Strongly Agree

24. I gave my goals (professional practice and student learning) a lot of thought when considering them.
    Strongly Disagree
    Disagree
    Agree
    Strongly Agree

25. The teacher evaluation process has caused me to make changes to the delivery of my curriculum.

Strongly Disagree
Disagree
Agree
Strongly Agree

26. I find myself creating more nontraditional assessments to ensure I am proficient in my evaluation.

Strongly Disagree
Disagree
Agree
Strongly Agree

27. I have become a more reflective teacher due to the new evaluation system.

Strongly Disagree
Disagree
Agree
Strongly Agree

28. I find myself collaborating with colleagues more since the implementation of the new evaluation system.

Strongly Disagree
Disagree
Agree
Strongly Agree

29. My curriculum already aligns with “proficient” evaluation; it was not necessary to adapt my delivery of it.

Strongly Disagree
Disagree
Agree
Strongly Agree

30. The self-reflection process, as prescribed in the Educator Evaluation model, has caused me to consider my teaching practice.

Strongly Disagree
Disagree
Agree
Strongly Agree

31. I find formative meetings with my evaluator important to my professional growth.
    Strongly Disagree
    Disagree
    Agree
    Strongly Agree

32. Since implementing the new evaluation system, I have created more rigorous standards-based units.
    Strongly Disagree
    Disagree
    Agree
    Strongly Agree

33. The Massachusetts Educator Evaluation system has not caused any change to my professional practice.
    Strongly Disagree
    Disagree
    Agree
    Strongly Agree

34. I prefer mini-observations as opposed to "traditional" full-period observations.
    Strongly Disagree
    Disagree
    Agree
    Strongly Agree

35. The new Educator Evaluation system is just another thing I have to do.
    Strongly Disagree
    Disagree
    Agree
    Strongly Agree

36. Teacher evaluation has caused me to rethink the way I assess students.
    Strongly Disagree
    Disagree
    Agree
    Strongly Agree

37. The evaluation process has resulted in positive changes in my professional practice.
    Strongly Disagree
    Disagree
    Agree
    Strongly Agree

38. I find myself making more creative lessons since the new Educator Evaluation system.
Strongly Disagree
Disagree
Agree
Strongly Agree

39. I care about my teacher evaluation rating.
Strongly Disagree
Disagree
Agree
Strongly Agree

40. The new Educator Evaluation system has caused me to think deeper about my lesson planning.
Strongly Disagree
Disagree
Agree
Strongly Agree

41. Teacher evaluation is a necessary evil.
Strongly Disagree
Disagree
Agree
Strongly Agree

42. I find summative meetings with my evaluator important to my professional growth.
Strongly Disagree
Disagree
Agree
Strongly Agree

43. Setting goals (professional practice and student learning) has helped me focus on improving my teaching practice.
Strongly Disagree
Disagree
Agree
Strongly Agree

44. I take seriously feedback from my evaluator about what they observe in my classroom.
Strongly Disagree
Disagree
Agree
Strongly Agree

45. Teacher evaluation has made me create lessons with more measurable outcomes.
Strongly Disagree
Disagree
Agree
Strongly Agree

46. I feel that my evaluator has a stake in my professional growth as an educator using the evaluation system.
Strongly Disagree
Disagree
Agree
Strongly Agree

47. The Massachusetts Educator Evaluation system has been implemented in a way that is fair.
Strongly Disagree
Disagree
Agree
Strongly Agree

48. The Massachusetts Educator Evaluation system has made me collaborate more with other educators.
Strongly Disagree
Disagree
Agree
Strongly Agree

Section 3: After completing this section, please click "submit" to enter your submissions. Thank you.

49. Your answers are completely anonymous, however, if you would like to enter your name in a drawing to win an Amazon gift card, please fill in your name and email below, and you will be contacted if you win. Your name will not be used for any reason other than for the purpose of picking a winner for this raffle.

50. If you have any additional comments about the Massachusetts Educator Evaluation system, please leave them in the comment box below. Any and all feedback is appreciated. Thank you for taking the time out of your day to complete this survey!