

# Promising Practices to Build Evaluator Capacity

## About This Brief

This brief describes promising strategies for relieving evaluator burden and improving consistency of evaluator practices under the Massachusetts Educator Evaluation Framework.

Since 2012, the Massachusetts Department of Elementary and Secondary Education has contracted with SRI International, Abt Associates, Nancy Brigham Associates, and J Koppich & Associates to study the implementation of the Educator Evaluation Framework. This Case Study Brief Series highlights promising strategies for implementing the framework in districts across the Commonwealth.

For more information on the study, please visit <http://www.doe.mass.edu/eval/resources/study/default.html>.

As part of a study of the implementation of the Massachusetts Educator Evaluation Framework (EEF), SRI International and its partners (Abt Associates, Nancy Brigham Associates, and J Koppich & Associates) visited Massachusetts school districts in spring 2015 to identify and share promising evaluation practices. The research team interviewed teachers, principals, and administrators in three districts—Revere, West Springfield, and Northbridge Public Schools—to learn about their promising strategies related to building evaluator capacity.<sup>1</sup> Evaluator capacity encompasses

- 1) evaluators' knowledge of and skills to conduct fair and comprehensive evaluations;
- 2) consistency among evaluators across schools, grades, and subjects in what they look for in observations, in how they provide feedback that improves instruction, and in how they rate educators in each performance level; and
- 3) time for evaluators to complete evaluation activities, including conducting observations, providing feedback, and reviewing evidence.

The research team specifically addressed these three districts' efforts to calibrate evaluators' feedback and ratings for school staff and to reduce evaluators' workloads associated with the new evaluation system.

Each of the three districts has been implementing the EEF for at least 3 years. Revere, an early-adopter district, began implementing in 2011–12. West Springfield and Northbridge, both Race to the Top districts, implemented the EEF in 2012–13.<sup>2</sup> West Springfield and Northbridge are Level 3 districts, and Revere is a Level 2 urban district.<sup>3</sup> Geographically, the three districts are located throughout Massachusetts, with one district each in the eastern, western, and central regions. Enrollments in the three districts range from approximately 2,500 to 7,000 students.

This brief outlines a series of promising strategies from the three districts for increasing consistency in evaluator practices and reducing evaluator workload, two issues with the EEF that continue to challenge districts across the Commonwealth. In fact, in a recent statewide survey of educators conducted as part of this study, 81 percent of staff agreed or strongly agreed that their evaluator's assessment of their own practice was fair, but 58 percent disagreed or strongly disagreed that educators were evaluated consistently across grades, subjects, and schools. Furthermore, 72 percent of principals disagreed or strongly disagreed that they had adequate time to evaluate teachers at their school. The brief begins with a description of a holistic approach for achieving evaluator consistency and then describes methods for relieving evaluator burden. The brief ends with a set of considerations for other districts that are looking for ways to improve the capacity of their evaluators to conduct fair and consistent evaluations under the new system.

1 The research team visited Revere and West Springfield Public Schools for 2 days each and conducted interviews and focus groups with district leaders, school principals, and teachers. The research team conducted phone interviews with educators in Northbridge Public Schools.

2 All other districts began implementing in 2013-14.

3 All Massachusetts districts and schools with sufficient data are classified into one of five accountability and assistance levels, with the highest performing in Level 1 and lowest performing in Level 5. In general, a district is classified into the level of its lowest performing school.

## A Holistic Approach to Evaluator Consistency

District administrators and school leaders in Revere, West Springfield, and Northbridge Public Schools recognize that establishing consistency across evaluators in their feedback and ratings is critical to ensuring that educators view the evaluation system as fair and valuable. Each of the districts has implemented a combination of strategies to address this issue, from bringing evaluators together to analyze evaluator feedback and teacher videos to providing content-specific training for evaluators. To complement district efforts, some school leaders have initiated their own efforts to ensure consistency among evaluators in their schools.

The one reason why I'm actually proud to be part of this district is we're never satisfied.... We'll never say that we've had enough calibration because we're always working to improve. — *School Leader*

### Promising Strategies to Promote Consistency Among Evaluators Across Schools

The three districts have provided opportunities for evaluators to become consistent in what they look for during observations and how they provide feedback. The efforts district administrators have initiated and implemented to calibrate evaluators across different schools are as follows.

**Analyze anonymized feedback and teacher videos.** Evaluators reported that collectively reviewing and discussing examples of feedback on real-world teaching support consistency. In Revere and West Springfield, evaluators met as a cohort to analyze anonymized feedback. District administrators randomly selected excerpts from evaluators' actual feedback to teachers. As a group, evaluators examined whether the feedback was useful and how it could be improved to spur changes in instructional practice and, by extension, student performance. In Revere, evaluators also met as a cohort to watch and rate videos of teaching. According to one district administrator, these activities helped “get everybody looking for the same things, observing things the same way.” Principals also found the process valuable for seeing how to frame feedback for teachers because, as one said, “It's easy to write something and it's a different thing to read it and really appreciate how it could be taken in many different ways. It was helpful to get some of those glaring examples.”

**Provide content-specific training.** Content-specific training helps equip evaluators to observe and give teachers feedback within specific subject areas. In Revere, curriculum directors facilitated sessions on giving feedback and observing math and English language arts teachers. Evaluators discussed “what looked good, what feedback [they could] give, what assumptions [they could] not make, which led to [discussions about] what is rigor and what that sounds like.” These trainings incorporated instruction on the new Curriculum Frameworks in addition to sessions for watching short teacher videos and practicing feedback.

**Conduct walk-throughs.** Revere and Northbridge promoted evaluator consistency by supporting small groups of evaluators in jointly conducting classroom walk-throughs. In Revere, each educator was assigned a primary evaluator and multiple secondary evaluators. These secondary evaluators could be another evaluator at the same school or other

I think that administrators need to see what other administrators are saying and have that conversation and decide what's fair and consistent. If teachers don't think things are fair, they're going to lose sight of the system.

—*Teacher*

relevant district administrators, such as curriculum directors. The group walk-throughs provided an opportunity for primary and secondary evaluators to compare their feedback and complement each other's knowledge of content or context. As one administrator reported, “We [bring] different strengths to the positions. When I do a walk-through with a vice principal, they know the students well. They might know [the] context of the student. When there are tricky decisions, typically we're in agreement.”

In Northbridge, a group of administrators, union representatives, and teachers convened monthly to conduct walk-throughs in small groups at a single school. These walk-throughs were conducted solely for evaluator training; they had no evaluative consequence for teachers. The school leaders who hosted the walk-throughs identified a focus area for the observations (e.g., student discourse). Teams observed three classrooms each and reconvened to debrief on their observations. A district administrator then provided the school leaders with a written summary of school-level feedback. One administrator said this process was valuable because of the emphasis on evaluator consistency, the specific school-level feedback for school leaders, and the transparency of the process with teachers.

**Set clear expectations.** Educators in Revere and Northbridge emphasized that both teachers and evaluators clearly understood their district's expectations for the format of evaluation feedback. This shared understanding was a direct result of reinforcing expectations during the ongoing training provided to evaluators described above. In Revere, all evaluators presented feedback by writing a narrative of what they observed in the classroom followed by a subjective

comment or probing question. Similarly, evaluators in Northbridge all used a systematic observation and feedback form that required the evaluator to include a question, commendation, and recommendation.

**Use technology.** Technology can potentially aid districts' calibration efforts, as evidenced by Revere, which created its own technology platform when the district began implementing the EEF. The platform houses all written evaluation activities, including evidence and feedback from unannounced and announced observations and summative evaluation reports, and includes a "communications window" in which evaluators and those being evaluated can engage in ongoing dialogue about observations. District administrators can access all educator evaluation files through this platform. They review evaluators' feedback to teachers approximately every 3 weeks and provide evaluators with assessments of their feedback to reinforce consistent practices across the district.

### Promising Strategies to Promote Consistency Among Evaluators Within Schools

Consistency among evaluators at the same school is as important as consistency among evaluators across schools. In Revere and West Springfield, a few principals took the initiative to meet with the other evaluators in their schools to support consistency efforts. This approach provided evaluators with additional opportunities to ensure consistency in how they provided teachers with feedback and ratings. At one school in West Springfield, for example, the principal met with her assistant principals daily to make sure "there is uniformity between the three of us... [and] come up with ways of formatting and explaining to people [why they received certain ratings]." The principal presented clear examples of exemplary work for each standard to both evaluators and educators so that all parties understood what was required for each rating level. This clear communication facilitated a high level of trust between educators and evaluators, leading educators to perceive evaluations within the school as "absolutely consistent." At a school in Revere, the principal helped guide a new assistant principal through the feedback process by providing examples of previous write-ups to ensure consistency between the two of them. Together, the principal and assistant principal also regularly conducted observations, talked through feedback, and wrote comments to educators.

We are a team here in every way, shape, and form. We have consensus. We need to have a common understanding of what it is [to be exemplary]. – School leader

A holistic approach, with varied strategies for promoting consistency initiated by both school and district leaders, provides an opportunity to calibrate the observation and feedback cycle districtwide. In addition to improving evaluator consistency, administrators will need to find ways to increase the amount of time that evaluators have to spend on evaluations.

### Relieving Evaluator Burden

In statewide surveys and interviews over the past 3 years for this study, evaluators have consistently identified the time demands of their evaluation responsibilities as a major challenge.<sup>4</sup> Providing high-quality feedback and analyzing multiple sources of evidence for each educator is a large undertaking. District administrators in Revere and Northbridge have implemented strategies to reduce the burden of the evaluation system on evaluators.

According to a recent statewide survey, 61 percent of principals are responsible for evaluating 21 to 40 educators.

**Leverage district administrators.** According to a statewide survey administered as part of this study, the majority of evaluators are responsible for evaluating at least 20 educators. In Revere, curriculum directors (e.g., in math, science, humanities, special education) serve as primary evaluators for a subset of school staff members who teach subjects aligned with their content area expertise. This strategy reduces the number of educators for whom evaluators at the larger schools are responsible. Educators in Revere found this approach to be effective because these curriculum directors interacted frequently with high school educators and provided them with valuable content-based feedback.

**Revisit district guidance on observation frequency.** Districts requiring frequent observations may be unintentionally contributing to increased evaluator burden under the EEF. Northbridge initially required evaluators to conduct 10 observations per year for both teachers with professional teaching status (PTS) and teachers without professional teaching status (non-PTS). To minimize the time burden for evaluators, district administrators established new guidance that reduces the number of observations required in a single year for school staff. The district now requires evaluators to observe non-PTS teachers six times per year and PTS teachers at least three times per year. An evaluator

<sup>4</sup> SRI International Research Brief: Implementation of the New Educator Evaluation System in Massachusetts, July 2014.

said the change in policy significantly relieved the burden of observations. A district administrator explained that the policy helped improve the overall quality of feedback in the district, which had been vague, generic, and unhelpful simply because evaluators lacked the time needed to craft useful feedback.

***Streamline evidence collection.*** In our interviews, educators consistently cited evidence collection as a time-consuming component of the new evaluation system for both evaluators and school staff. In response, West Springfield revised its evidence collection process. Educators now compile a “directory of evidence” in which they list and briefly describe the artifacts related to their progress on meeting rubric standards and indicators. Educators are expected to produce the actual artifacts only if their evaluators request them. This process has reduced the time it takes educators to collect evidence and the time evaluators spend reviewing teacher-generated artifacts.

In 2014–15, Northbridge’s educator evaluation committee (a team of union representatives, teachers, and administrators) met to discuss the key aspects of strong classroom instruction. After a few meetings in fall 2014, the committee identified a prioritized set of indicators from the educator performance rubrics.<sup>5</sup> District administrators revamped guidelines for evidence collection, which had previously required that teachers submit two to three pieces of evidence for every indicator in the rubric and placed a large time burden both on educators to compile evidence and on evaluators to assess the evidence. Under the new guidance, educators are required to submit evidence for four key elements.<sup>6</sup> An administrator explained that this minimizes burden on teachers, as well as principals. She stated, “The evidence collection is way more streamlined. Now, I get folders [instead of binders].” Furthermore, the streamlined guidance results in a more meaningful and thoughtful process for educators. An administrator explained that her teachers were much more thoughtful in the evidence collection process; she noted, “I felt like it helped them [teachers] to be very reflective, and that’s something I really want my teachers to be.”

---

The three districts developed promising practices to reduce evaluator workload, but they acknowledged that this work is far from complete. Interviewees suggested further strategies for relieving burden, such as peer evaluation and reducing the frequency of evaluations for veteran teachers with a history of positive reviews, but districts in the Commonwealth have not yet implemented them.

## Considerations for Other Districts

Under the new evaluation system, evaluators play a key role in supporting instructional improvement. Revere, West Springfield, and Northbridge have implemented promising strategies to promote evaluator consistency within and across schools and to relieve burden on evaluators. Findings from these districts’ efforts to support evaluators might be instructive for other districts.

- Districts should provide varied and ongoing opportunities for evaluators across schools to become consistent in their understanding of the performance rating levels and in their manner of providing feedback. These efforts should be supplemented with formal and informal collaboration among evaluators at the same schools. Implementing this holistic approach helps ensure that all educators receive clear communications about evaluator expectations for feedback and timing of observations and ultimately heightens educators’ perceptions of the system’s fairness overall.
- Districts can use technology in a variety of ways to increase consistency, enhance feedback, and customize platforms to highlight their expectations for the evaluation cycle. When used intentionally, technology can provide a feedback loop between administrators and evaluators to reaffirm expectations for evaluators set by the district.
- Districts have more work to do to provide evaluators with more time to provide meaningful feedback on educators’ instructional practices. Leveraging district administrators as additional evaluators and adjusting guidelines for the quantity of observations and evidence collection are steps in the right direction. However, it will be useful to find additional ways to decrease the number of educators evaluators are responsible for or to expand the number of individuals conducting evaluations.

---

<sup>5</sup> “In the ESE Model Rubrics, descriptions of practice associated with the Standards and Indicators of Effective Teaching Practice are broken down into elements and distributed across four performance levels (Unsatisfactory, Needs Improvement, Proficient, and Exemplary). Each element describes a discrete behavior or related set of behaviors that educators and evaluators can prioritize for evidence-gathering, feedback, and eventually, evaluation.” Source: Massachusetts Educator Evaluation: Rating Educator Performance, April 2013.

<sup>6</sup> An evaluator can request additional evidence from teachers if needed.