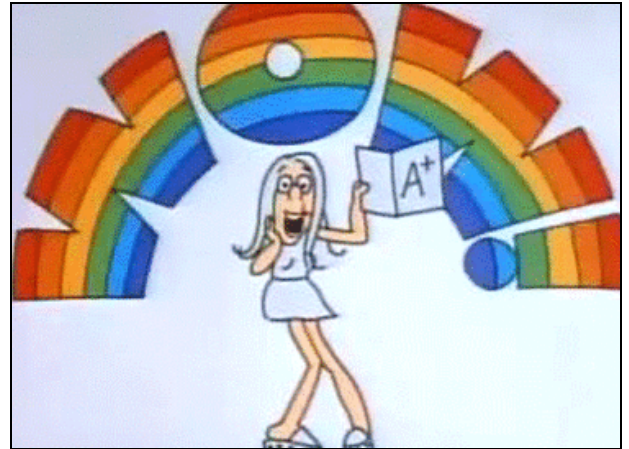


Peer to Peer: Exemplary Ratings

By Mark Lonergan with Anne Slater

It was May 2001, at the end of my first year of teaching in Boston Public Schools, when I was pulled into the headmaster's office to go over my end-of-year evaluation. Sensing that I was nervous, my headmaster started with some compliments and said that I'd had a very strong first year. I was happy to have almost made it to June and relieved that my 2nd period class had slowly found a way to work and learn together. Near the end of our meeting, she handed me my formal BPS evaluation. I was a bit surprised to see my final rating: "Meets or Exceeds Standards." *What's this?* I thought, *Didn't you just say that I was doing a good job? Isn't this the equivalent of telling a student, "You're an A, B, or C student"?*

In 2011, with the passage of State Regulation 603 CMR 35.00, the rating of "Meets or Exceeds Standards" changed in Boston and across the state to our four current rating levels. While policymakers and administrators have spent lots of time calibrating and thinking about how to distinguish between unsatisfactory, needs improvement, and proficient ratings, there's been much less consensus around trying to distinguish between proficient and exemplary--in trying to articulate the difference between an "A" or "B" or "C" teacher. This month, we'll take a closer look at exemplary ratings: What are they? How are they determined? And what parts of the process might need to change in the future?



SCHOOL and DISTRICT DATA

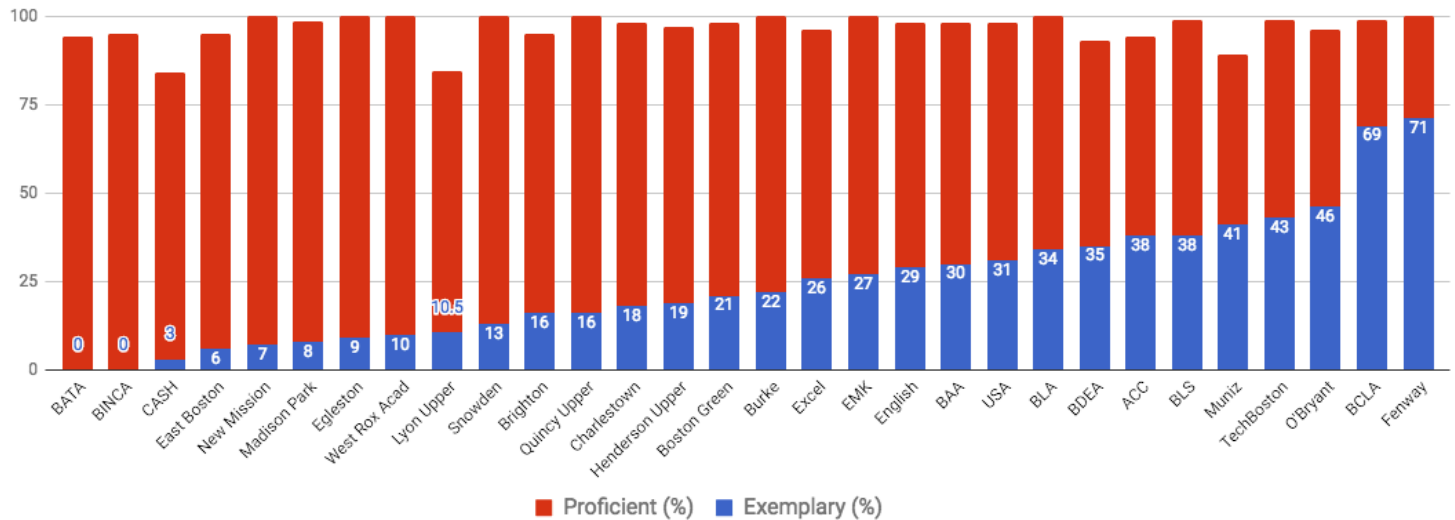
According to the "School and District" profile data available on the DESE website, about 11.5 percent of educators across the state were awarded an exemplary rating in 2016 (the most recent data available to the public). By comparison, 84 percent are proficient, 4 percent are needs Improvement and 0.5 percent are unsatisfactory statewide. In Boston, about 20 percent of BPS's 3800 teachers were exemplary in 2016.

One challenge that emerges in looking at the DESE data is how widely the share of exemplary ratings can vary from district to district and even within districts. At the Massachusetts Academy for Math and Science in Worcester, 8 of 10 staff members were rated as exemplary. But some districts (including Arlington) had no teachers at all with an exemplary rating. Claire Abbott, the Educator Effectiveness Coordinator at the Department of Elementary and Secondary Education, said that there's no quota or target determined by the state. "From the start, we've wanted the educator evaluation framework to remain responsive to the local context of individual districts and to the professional judgment of evaluators within those communities, so we have no recommendation for how many educators should be rated as exemplary." The towns of Orleans, Wellfleet, Provincetown, Brewster and Eastham all had more than 50 percent of teachers rated as exemplary. While Arlington, along with Bedford, Lexington, Marblehead, Richmond, Rockland, Northborough and Harvard, all had zero exemplary teachers in 2016.

The frequency of exemplary ratings also varies widely within BPS. The chart shows 2016 data for Boston high schools. Some schools had more than half of the staff rated exemplary, and some schools had none. This is because Boston, like the state, leaves the determination of ratings up to individual evaluators. Jerome Doherty, director of performance management for Boston said, "There are no formulas for arriving at an overall rating (aside from the requirement that Standards I and II must be rated at least proficient in order to issue an overall proficient rating), so we advise using sound reasoning and common sense to arrive at an appropriate rating that is representative of the educator's overall practice as a whole."

BPS High School Exemp & Prof Ratings 2016

From DESE "School and District Profiles"



This may mean that different evaluators have a different understanding of what exemplary should look like. Carmen Davis, principal of Irving Middle School said that she gives out exemplary ratings very rarely. "In terms of exemplary ratings, a teacher should be able to implement effective practice and also understand the why behind what they're doing," said Davis. "They should understand what makes the practice effective related to impact on student learning and be able to teach their colleagues about their instructional practices."

Elis Kanner, who is the director of coaching and evaluation at Charlestown High School, said that there are three main problems with the current expectations for exemplary ratings. "Right now, the system isn't transparent, teachers don't know in advance what they're getting so there's no agency, and there's inconsistency from building to building and from evaluator to evaluator."

CALIBRATION CHALLENGES

Laura Chesson, a former BPS teacher who is now the Superintendent of Groton-Dunstable, said "In my former district, Arlington, we spent several meetings hammering out a specific definition for exemplary, and determining how a teacher got that rating. The teacher had to indicate that he/she wanted to work towards that rating, with a heavy emphasis on teacher leadership roles." In Groton-Dunstable, Chesson and Assistant Superintendent Katie Novak have created a menu to help teachers understand exactly what is expected for those wanting an exemplary rating. "This is a question we've been struggling with for years with our educator evaluation working group," said Novak. Their menu of exemplary practices tries to clearly define what it means to "model for others." In Groton-Dunstable, educators can be a model by presenting at a staff meeting, leading a department, sharing expertise at a conference or seminar, or by working as a peer coach/mentor.

In Boston, the variance between schools and evaluators is wide. One challenge is that the language within the exemplary ratings leaves room for interpretation. Most standards include the phrase "is able to be a model" to distinguish between proficient and exemplary ratings. Does "able to model" mean that a teacher is willing to model or has actually been a model? Tony King, headmaster of Boston International and Newcomers Academy, feels that "able to model" is best demonstrated by working with fellow staff members and maybe leading PD. "The issue that had been holding us back was the part of exemplary that suggests the teacher is able to model the standard," said King. "We didn't have enough opportunities for teachers to do this, but we have added more teacher-led PD in the last two years, so I anticipate a solid number of exemplary rated teachers at BINCA."

DESE's Claire Abbott said that the variance in Boston "aligns with the data we have in general. There is variation in terms of the use of the exemplary rating across districts." Abbott says that the variation is higher within districts than across districts. "I'm not surprised to hear that there's school-to-school differences." This variation is one of the primary motivators for Abbott's current project: an update of the 2011 rubric which will include refinements to modeling language in the exemplary descriptors.

BTU Vice President Erik Berg said he is also not surprised by the variance around exemplary ratings. “The personalities behind the use of any rubric are inherently subjective,” said Berg, “So no matter how calibrated and defined a system is, the inevitable differences across human judgment appear when indicators like ‘all of the time’ versus ‘most of the time’ are in play.”

TRANSPARENCY

Another challenge with the current exemplary rating is that some people say that it’s not really clear how exemplary is defined. Jess Madden-Fuoco, who is the director of teaching and learning at Boston Green Academy and who also teaches the BPS “Observation and Feedback” courses, said that she feels the rubric definitions are pretty clear. “I can clearly see a difference between proficient and exemplary, but there is a wide range of how evaluators within schools and across the district assign proficient vs. exemplary ratings,” said Madden-Fuoco. “It’s frustrating when the *Globe* runs stories that make it sound like teaching and learning is stronger in buildings with higher rates of exemplary ratings when the truth could be that evaluators in those buildings are more likely to rate educators as exemplary.

The Office of Human Capital is trying to make criteria clearer across the district. OHC’s Jerome Doherty said, “We are working on developing guidance for how school leadership should think about the evidence in comparison to the exemplary descriptors on the rubric. We want to avoid school administration saying ‘you must do X to be rated exemplary,’ where X represents a very specific teaching practice or artifact.” Doherty and the Office of Human Capital are gathering examples of how schools across the district are addressing the question of exemplary ratings.

MORALE

Another challenge with exemplary ratings is that they can have a negative impact on morale, for those teachers who don’t receive an exemplary rating and even for those who do. “Morale in BPS can already feel tough. We get love from students and hopefully from colleagues, but as a system it can be tough to feel appreciated,” said Madden-Fuoco. “This proficient vs. exemplary issue can be demoralizing in an unintended way, especially for people who are aiming for exemplary but do not get there.”

Madden-Fuoco also said that once a teacher has reached an exemplary rating, it can also be challenging to adjust to anything less than exemplary. “People can start to see proficient as not a great rating. If you read the language in the “proficient” category of the rubric, it is not negative, but proficient can feel like a lower rating because it isn’t exemplary. One teacher [who moved from exemplary to a proficient rating] told me it felt like a fall from grace,” she said. Irving principal Carmen Davis agreed: “It can be hard when evaluation ratings change. It would be good if people were more normed.”

Veteran educator Mary O’Brien of the Henderson K-12 school said that she tries not to worry too much about exemplary ratings. “Advice I got early in the process was that we should aim for exemplary but be happy with proficient ratings,” said O’Brien. “When this was first rolled out, I remember someone saying that exemplary should be hard to get and will be difficult to maintain. I’ve tried to remember that in my own practice. All of us can be learning.”

Ben Helfat, headmaster of Boston Adult Technical Academy (BATA), said that another challenge with exemplary ratings is that it doesn’t fit into a coaching or growth mindset. “If we’re going to adopt a coaching model,” said Helfat, “Then exemplary really shouldn’t be there. I don’t see where exemplary fits into a growth mindset.”

Helfat said that in his school he certainly has teachers who are doing exemplary work (and earning exemplary ratings on standards), but he said that “When a teacher in my school does something exemplary, I try to highlight it and might even email it to the rest of the staff, so that they can know about it and try to emulate it. Giving an ‘E’ rating privately to someone doesn’t have the same impact on the community.”

At Charlestown High School, Elis Kanner has been trying to provide coaching, support, and acknowledgement through their school-wide coaching and support program. Kanner and other evaluators at Charlestown are in the second year of piloting a school-wide coaching program that is also tied to EDFs and the evaluation process. She said that response from staff has been positive overall, but the “wrinkle of exemplary” has been a challenge for time-crunched teachers and administrators. “Teachers felt attached to having the opportunity to be rated exemplary, but we had to invent a fair process. I ended up spending more hours on the question of exemplary than most other parts of my job.” Kanner said that even after drafting a transparent process, it still felt like there was not enough time for evaluators to review, calibrate, and evaluate artifacts. “There’s not enough time to coach, let alone time to calibrate on what exemplary looks like,” said Kanner. “This was time away from being in teachers’ classrooms and meeting with teachers to actually work with them on improving their practice.”

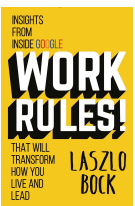
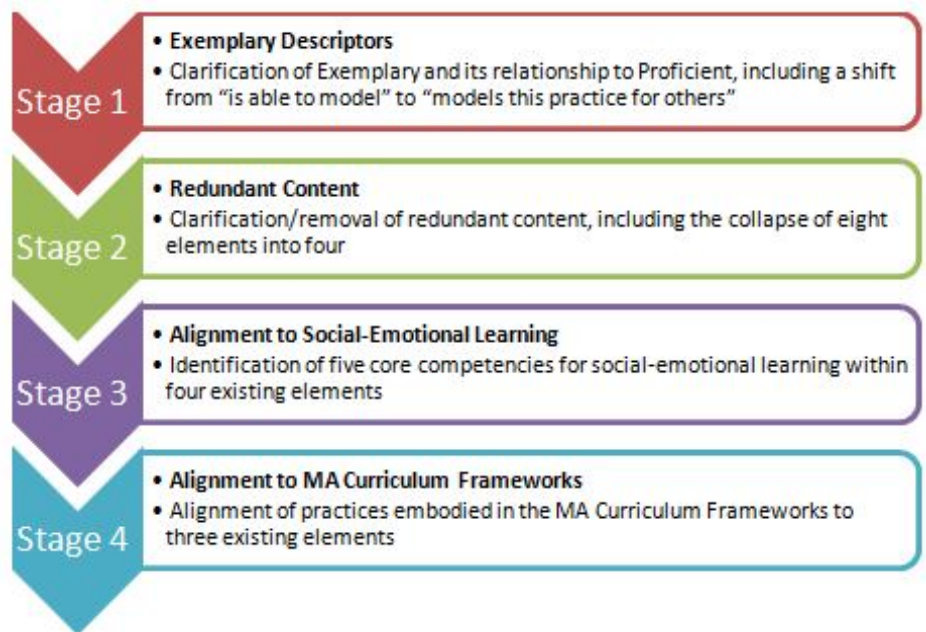
NEXT STEPS

Jerome Doherty hopes the district can begin to provide more clarity around the expectations for exemplary ratings. “At the district level, while we have developed clarifying language around our shared vision of effective practice through the Essentials for Instructional Equity, we should have a similarly transparent set of expectations for school leadership, which I am happy to say we are working toward at this very moment,” said Doherty. He also said that his office is hoping to develop “clearer and more user-friendly rubric language.” Doherty and OHC are also offering “guiding principles” for exemplary ratings that give school leaders a menu and options to think about how to define exemplary practice, in addition to updating the district’s interactive rubric (bostonpublicschools.org/ir).

Others think that we should not really be worrying about exemplary ratings at all. “I do not see any use or purpose behind having a distinction between proficient and exemplary in our evaluation system,” said Erik Berg. “I would eliminate the exemplary category. It only serves to sow discontent and allow unconscious biases to come to the fore.”

At the state level, Claire Abbott is in the middle of an update to the DESE 2011 rubric. “This rubric has been used for 6 years by over 85 percent of districts across the state. It’s now time to update it.” The proposed refinements include merging some rubric elements to streamline the rubric and eliminate some redundancies. For example, “Well Structured Lessons” (I.A.4) and “Standards-Based Unit Design” (I.A.3) are combined into one element, “Well-Structured Unit & Lesson Design,” in the updated rubric.

Another proposed change for the next iteration of the DESE model rubric will be a clearer distinction between proficient and exemplary practice. Abbott said that we “want anyone who’s looking at the two ratings side-by-side to clearly see how to move from proficient to exemplary.” The committee of teachers, administrators, and policy-makers who are revising the rubric recommended changing language from “is able to model” to “models this practice for others.” The advisory board also recommended removing the “modeling” language from 7 rubric elements, like “Sharing Conclusions with Colleagues” (I.C.2). All of the proposed refinements are being piloted in seven districts across Massachusetts, with final updates to be determined based on their feedback by July.



The challenge of recognizing exemplary performance is not unique to the teaching profession. In fact, it may be even more of a challenge in other career fields, where performance evaluation is often tied to how much you are paid or how much of a quarterly bonus you’re awarded. In the book *Work Rules*, Google’s vice president of people operations Laszlo Bock explains that there are five different performance levels at Google: needs improvement, consistently meets expectations, exceeds expectations, strongly exceeds expectations, and superb. Perhaps distinguishing between “strongly exceeds expectations” and “superb” is even more challenging than distinguishing between “proficient” and “exemplary.”

But for now, Elis Kanner is worried that exemplary ratings are not actually helping our students. “I don’t know how exemplary ratings actually helps kids. I tell teachers that if you want recognition and challenge, you should pursue National Board certification instead. Ideally I hope we can find other ways to honor teacher practice.”

The Peer-to-Peer column is written by Anne Slater, from the Peer Assistance program, and Mark Lonergan, from the Peer Assistance and Review (PAR) program. This column is written by teachers and for teachers. If you have a topic you’d like us to explore, please email us. To find out more about what Anne and Mark and others do as peer assistants and consulting teachers, visit btu.org/whats-working/peer-mentoring/ or bostonpar.org.