

How to Make Teacher Evaluations Accurate, Fair, and Consistent

By Kim Marshall

New teacher-evaluation policies are being implemented pretty much everywhere with rubrics, more-frequent classroom visits, student surveys, and value-added test data. As the stakes rise, teachers, principals, superintendents, and school boards have a lot to worry about. How should principals document what they see in classrooms? How and when should they use rubrics? What role should student achievement play?

Concerns like these explain why so many principals are being sent to off-site training to improve what's known as inter-rater reliability. Participants take notes on classroom videotapes, give rubric scores, compare their evaluations with the "correct" answers, and are (usually) declared "certified." This kind of training can be a helpful exercise, developing principals' sense of what good (and not-so-good) teaching looks like and clarifying what's most important in classrooms.

But if superintendents want teacher evaluations to be accurate, fair, and consistent, they need to address five key issues within their districts: (a) getting principals to make enough classroom visits to see daily reality; (b) ensuring that every principal really does have a good eye for instruction; (c) polishing principals' skills at giving feedback to teachers; (d) deciding how and when to use the district's rubric; and (e) keeping student learning at the center of supervisory conversations. Here are some suggestions for reaching that goal:

- **Classroom visits.** Let's face it: Teacher evaluation based on infrequent, announced classroom visits is inaccurate, ineffective, and dishonest. To see how teachers are performing day to day, principals need to be in each classroom at least once a month for short, unannounced visits, followed by face-to-face conversations. Superintendents should hold principals accountable for knowing their teachers and continuously coaching (and praising) them.

- **Observation skills.** For starters, it's essential that all principals commit to giving honest feedback when they see mediocre and ineffective teaching (as defined by the district's rubric). How can superintendents be sure of principals' judgment and courage? By visiting each school on a regular basis, observing a couple of classrooms with the principal, stepping into the corridor, and having the principal identify two or three key teaching points. All principals need this kind of boots-on-the-ground supervision (and affirmation). Of course, for superintendents to make regular classroom visits, they must have a manageable span of control. In large districts, this means creating clusters of no more than 12 to 15 schools, each

supervised by an area superintendent. (Boston and Newark, N.J., have recently reorganized along these lines.)

Superintendents can also check on principals' classroom-observation skills by conducting online surveys of teachers with questions like: Do you have confidence in your supervisor's knowledge of your subject or grade? Has your principal's feedback on your teaching been helpful?

• **Feedback skills.** Principals' classroom visits will accomplish very little if they don't talk to teachers afterward (and then send brief narrative write-ups). "Let's face it, teacher evaluation based on infrequent, announced classroom visits is inaccurate, ineffective, and dishonest."

How can superintendents monitor these interactions? First, by role-playing with principals right after each classroom visit, with the superintendent playing the teacher. Second, by occasionally sitting in on principals' feedback conversations with teachers. Third, by making sure principals co-observe with instructional coaches to build their content knowledge (just because students are sitting around a kidney-shaped table reading a nice book with their teacher doesn't mean that high-quality guided reading is going on).

Fourth, superintendents should read a sampling of principals' classroom write-ups. And finally, superintendents should devote a segment of their monthly principals' meetings to this sequence:

- Play a 10- to 15-minute classroom videotape and give principals a minute to look over the notes they have taken.
- Have principals pair up and role-play the principal-teacher conversation.
- Have the "teacher" give candid feedback on how the "principal" handled substance, choice of words, eye contact, and body language.
- Have principals switch partners and switch roles (so everyone gets to be the principal, but with a different partner) and do the role-play and debrief again.
- As a whole group, discuss the most important points, tactics for presenting them to the teacher, and ways to handle push-back.
- Have principals write a paragraph to the videotape's teacher summing up the feedback and the conversation, and then read theirs aloud to another principal. Share one or two write-ups with the whole group.

This 60- to 70-minute process puts everyone on a steep learning curve developing the skills, confidence, and courage to do this work.

• **The rubric.** Trying to fill out checklists or rubrics during classroom observations prevents principals from being good observers. They should be looking over students' shoulders, asking one or two students, "What are you working on today?"; scanning wall displays; and listening intently to classroom interactions and jotting a few notes. Nor should rubrics be used immediately after leaving the classroom; that's when administrators should be deciding on the one or two most important pieces of feedback.

Instead, rubrics should be used at three strategic points during the year: In September, teachers self-assess and agree on two or three personal-improvement

goals with their supervisors. At midyear, principals check in with each teacher to see if he or she is in general agreement on rubric scores (each fills out the rubric before the meeting, and they compare ratings page by page and debate any differences). At the end of the year, they repeat the midyear process, again discussing any disagreements. The principal decides on summative ratings.

A common misconception is that principals need to gather evidence on every single rubric element. If school leaders were required to write their end-of-year teacher evaluations from memory, copious evidence-gathering would make sense. But a good rubric is like a multiple-choice test: Its detailed four-level descriptions of teaching behaviors act as a memory prompt, pulling out information about each teacher's performance through the year from multiple observations, interactions, and brief write-ups. Any gaps or misperceptions are filled in by the teacher's self-assessments. This is an amazingly efficient way to review the year's performance, usually taking less than half an hour per teacher.

What about the tendency toward grade inflation in teacher evaluations, so common in the past? Superintendents should compare teachers' rubric scores with classroom observations during school visits ("You gave her a 'highly effective' in classroom management? That seems too generous!"). Trying to fill out rubrics after isolated classroom videotapes in principal meetings with superintendents is less helpful, since it's impossible to give comprehensive scores on a 10-minute classroom clip, and single videotapes don't simulate the task of summing up a year's observations and interactions. Better for principals to see excerpts from a videotape portraying a full year of instruction and then compare scores on selected rubric components.

• **Student learning.** How can districts deal with discrepancies between teachers' rubric scores and their students' achievement? The worst strategy is to wait for end-of-year test scores, which don't arrive until summer. Throughout each year, superintendents need to be in schools asking: Are teachers checking for understanding as they teach and then immediately putting students' responses to work? Are principals monitoring teacher teams as they analyze and follow up on unit tests, interim assessments, and performance tasks? Do the results of in-school assessments jibe with teachers' ratings?

Perfect inter-rater reliability is unattainable; schools are way too complex. But if superintendents are in classrooms every week, surveying teachers, looking at other data, and using their monthly leadership meetings well, principals will up their game and teachers' evaluations will be increasingly accurate, fair, helpful, and consistent. And that will make a major difference to the quality of teaching and learning.

Kim Marshall coaches principals, consults, and speaks on school leadership and evaluation, and publishes the weekly online "Marshall Memo." Based in Massachusetts, he is a former public school teacher and administrator.

