

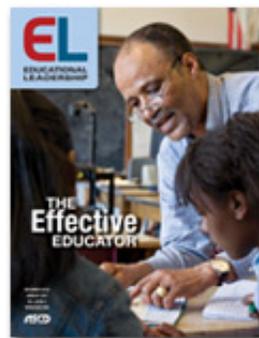
Evaluations That Help Teachers Learn

Charlotte Danielson

A good system of teacher evaluation must answer four questions: How good is good enough? Good enough at what? How do we know? and Who should decide?

"This is so much *better!*" commented Carla, a 4th grade teacher, following an evaluation conference with her supervisor:

Before, I had no idea what my principal was looking for—I had to be a mind reader! So I just played it safe, taught a familiar lesson, one I knew would go well—but did the process improve my teaching? Not at all! In my old school, the principal just came in with a checklist, but we never really talked. But this time, we had a great conversation about how to help my students *want* to write. It really made me think. As a result, I've got a new approach: I'm going to engage some students around the things they're passionate about and have them try to convince their classmates about the value of such interests.



The Problem

Carla's statement provides an insight into how we might improve teacher evaluation to better foster conditions for both teacher and student learning. Let's consider the deficiencies of traditional systems. These include

- Outmoded evaluative criteria, usually in the form of checklists.
- Simplistic evaluative comments, such as "needs improvement," "satisfactory," and "outstanding" without any consistency as to what those words mean. Many teachers end up being rated at the highest level on every item, with no guidance as to where they might focus their improvement efforts.
- The same procedures for both novice teachers and career professionals— no differentiation that reflects veteran teachers' experience and expertise.
- Lack of consistency among evaluators; a teacher might be rated at the highest level by one administrator and much lower by another. This makes it much easier to attain tenure in some schools than in others, a violation of a fundamental principle of equity.
- One-way, top-down communication. Evaluation is a process that's "done to" teachers, and it often feels punitive, like a "gotcha."

Why Do We Evaluate Teachers?

We can remedy these problematic characteristics by attending to some basic principles of assessment and teacher learning. First, it helps to be clear about why we even *have* teacher evaluation. Laws, of course, require it. But why are there laws? The first and most fundamental reason is because public schools are public institutions; they take public money, and the public has a right to expect high-quality teaching. But there are two more basic purposes.

To Ensure Teacher Quality

Credibility in an evaluation system is essential. A principal or a superintendent must be able to say to the school board and the public, "Everyone who teaches here is good— and here's how I know." A teacher evaluation system that satisfies this requirement will include the following:

- *A consistent definition of good teaching.* To assess the quality of teaching practice, it's essential to *define* it. It's not sufficient to say, "I can't define good teaching, but I know it when I see it."

One of the most widely used systems that define good teaching is the Framework for Teaching¹, which describes not only the teaching that occurs in the classroom but also the behind-the-scenes work of planning and other professional work, such as communicating with families and participating in a professional community. For each component of good teaching, the framework includes four levels of performance—unsatisfactory, basic, proficient, and distinguished—that describe the degrees of teacher expertise in that component. (See fig. 1 for the four levels of performance in questioning and discussion techniques.)

Figure 1. Levels of Performance in Questioning and Discussion Techniques

Element	Unsatisfactory	Basic	Proficient	Distinguished
Quality of Questions	Teacher's questions are virtually all of poor quality, with low cognitive challenge and a single correct response. They are also asked in rapid succession.	Teacher's questions are a combination of low and high quality, posed in rapid succession. Only some invite a thoughtful response.	Most of teacher's questions are of high quality. Adequate time is provided for students to respond.	Teacher's questions are of uniformly high quality, with adequate time for students to respond. Students formulate many questions.
Discussion Techniques	Interaction between teacher and students is predominantly recitation style, with the teacher mediating all questions and answers.	Teacher makes some attempt to engage students in genuine discussion rather than recitation, with uneven results.	Teacher creates a genuine discussion among students, stepping aside when appropriate.	Students assume considerable responsibility for the success of the discussion, initiating topics and making unsolicited contributions.
Student Participation	A few students dominate the discussion.	Teacher attempts to engage all students in the discussion, but with only limited success.	Teacher successfully engages all students in the discussion.	Students themselves ensure that all voices are heard in the discussion.

- *A shared understanding of this definition.* Everyone in the system—teachers, mentors, coaches, and supervisors— must possess a shared understanding of this definition. Having a common language to describe practice increases the value of the conversations that ensue from classroom observations. For example, discussing "student engagement in learning" is more effective when everyone understands what this looks like in light of four elements: activities and assignments, grouping of students, instructional materials and resources, and structure and pacing. Conversations using this more specific language invite teachers to analyze their own practice and invite observers to inquire about the decisions a teacher has made in planning and executing a lesson.
- *Skilled evaluators.* Those who support teachers—mentors, coaches, supervisors, and so on—must be able to recognize classroom examples of the different components of practice, interpret that evidence against specific levels of performance, and engage teachers in productive conversations about their practice. Evaluators must be able to assess teachers accurately so teachers accept the judgments as valid and the public has confidence in the results. Evaluations that focus on quality assurance yield judgments that are fair, reliable, and valid. They are helpful in looking at both new and experienced teachers' practice and in determining whether a teacher's skill has slipped below standard and needs strengthening. Administrators may use the evaluations for decisions regarding employment and compensation. This is crucial when deciding which teachers should attain permanent status as tenured professionals or which teachers should be nominated for leadership positions as mentors or coaches.

To Promote Professional Development

But there's another purpose of teacher evaluation: to promote professional learning. Teacher evaluation typically serves this more developmental purpose through professional conversations between teachers and colleagues who observe in their classrooms and between teachers and supervisors following formal or informal observations.

A commitment to professional learning is important, not because teaching is of poor quality and must be "fixed," but rather because teaching is so *hard* that we can always improve it. No matter how good a lesson is, we can always make it better. Just as in other professions, every teacher has the responsibility to be involved in a career-long quest to improve practice.

Two in One

The challenge is merging these two purposes of teacher evaluation. Educators need to create procedures that yield valid and reliable results—that is, that satisfy the legitimate demands for quality assurance while promoting professional learning. In truth, the demands are somewhat different. A system to ensure quality must be valid, reliable, and defensible—these are "hardsounding" qualities—whereas a system designed to promote professional learning is likely to be collegial and collaborative—these are much "softer-sounding" qualities.

Until recently, educators' attempts at merging quality assurance with professional learning have taken the form of enhancing evaluators' skills using techniques like clinical supervision and cognitive coaching. These are valuable skills and worth learning, but they are insufficient. The profession is better served when the requirements for these two purposes are embedded in the *design* of the systems themselves.

Getting from Here to There

We can get a clue as to the nature of this problem if we consider the typical observation, supervision, and evaluation process in use in most schools. The scenario proceeds as follows: The administrator goes to the classroom and watches a lesson, takes notes, goes away and writes up the notes, and then returns

and tells the teacher about the lesson (what was good, what the teacher could improve). Most observations are a variation on this theme.

It's important to note that in this scenario, the administrator is doing all the work; the teacher is completely passive. (The teacher has, of course, taught the lesson, but the teacher contributes nothing to the observation itself.) So it's not surprising that teachers don't find the process valuable or supportive of their learning. The process violates everything we know about learning—that learning is done by the *learner* through a process of active intellectual engagement.

If we want teacher evaluation systems that teachers find meaningful and from which they can learn, we must use processes that not only are rigorous, valid, and reliable, but also engage teachers in those activities that promote learning—namely self-assessment, reflection on practice, and professional conversation.

We can modify the traditional observation scenario to accomplish these aims. A revised process—like the one Carla was so enthusiastic about at the beginning of this article—might look like this:

1. The administrator goes to the classroom, watches a lesson, and takes notes on all aspects of the lesson: what the teacher says and does, what the students say and do, the appearance of the classroom, and so on.
2. The administrator gives a copy of his or her notes to the teacher.
3. The administrator analyzes the notes against the evaluative criteria and levels of performance.
4. The teacher reflects on the lesson using the observer's notes and assesses the lesson against the evaluative criteria and levels of performance. The teacher will probably, as result of this reflection, identify aspects of his or her teaching to strengthen, and that teacher will reach these conclusions without prompting from the principal. Of course, the principal can always point things out, but when the teacher reflects on a lesson before the post-observation conference, he or she will frequently be as critical as the principal would have been.
5. The teacher and the administrator discuss the lesson. The teacher puts the lesson into context for the administrator, and together they decide on the teacher's strengths and areas for growth. Naturally, the administrator wasn't in the classroom the previous day and can't be familiar with all the issues that the teacher must address. So the teacher might describe a particular student's learning challenges, and the principal might suggest a different approach. But they conduct the conversation in light of their shared understanding of what constitutes good teaching.

Seeing Benefits in Chicago

A recently published study of a two-year pilot program in Chicago Public Schools has documented the benefits of this approach. Conducted by researchers at the Consortium on Chicago School Research at the University of Chicago, the Excellence in Teaching Project aims to accurately measure a teacher's classroom performance. The project is the proposed replacement for a teacher evaluation checklist that administrators have used in Chicago Public Schools for the past 30 years. Principals in the pilot used the Framework for Teaching to guide their classroom observations as well as the required pre-observation and post-observation conferences.

The pilot provides an insight into the perceptions of participating teachers and administrators, who signaled the following areas as being crucial to effective teacher evaluation:

- *A consistent definition of good teaching.* For a teacher evaluation system to be transparent and credible, everyone—both teachers and administrators— must understand what constitutes good practice. Unless principals participate in focused

training, they probably will not have this understanding. But they appreciate acquiring that knowledge. As one principal noted, "The thing I like about the framework is that it actually makes you cognizant of the behaviors that constitute excellence in teaching."

- *Opportunities to engage in meaningful conversations about practice.* As members of the Danielson Group have observed after working with teachers and administrators in hundreds of school districts to enhance professional practices, "It's all about the conversation." Noted one teacher, "You get to close the door, turn off the noise, and actually sit and talk [with your supervisor], which is really, really nice."
- *A focus on what matters.* Both teachers and administrators appreciate an opportunity to concentrate their collective attention on the important issues of teaching and learning. These typically occur in the post-observation (reflection) conference. As one principal pointed out, "The conversation is entirely different. My conversation before was 'You were tardy,' 'You didn't turn in your lesson plans,' all those kinds of things. Now this conversation is about good instruction."

Two Challenges

The Need for Trained Evaluators

A credible system of teacher evaluation requires higher levels of proficiency of evaluators than the old checklist, "drive-by" observation model. Evaluators need to be able to assess accurately, provide meaningful feedback, and engage teachers in productive conversations about practice.

In our experience with the Framework for Teaching, members of the Danielson Group have trained hundreds of observers all across the United States and in other countries as well. Our findings have been somewhat humbling; even after training, most observers require multiple opportunities to practice using the framework effectively and to calibrate their judgments with others.

Most administrator preparation programs don't teach such skills; administrators must acquire them on the job. But when they do learn them, administrators can be the instructional leaders that schools so urgently need.

A training program for evaluators—one that uses the Framework for Teaching—consists of several steps:

1. Participants familiarize themselves with the structure of the Framework for Teaching, which consists of four domains of teaching responsibility (planning and preparation, classroom environment, instruction, and professional responsibilities); 22 components that describe those domains; and two to five elements that fully describe each component.
2. Participants learn how to recognize the sources of evidence for each component and element. For example, Domain 2 (the classroom environment) and Domain 3 (instruction) are demonstrated primarily in the classroom, whereas Domain 1 (planning and preparation) and Domain 4 (professional responsibilities) depend on artifacts, such as teachers' techniques for communicating with families (for example, newsletters or handouts for back-to-school night) or logs of professional development activities.
3. Participants learn how to interpret the evidence against the rubrics for each component's levels of performance. For example, in assessing whether a classroom creates an environment of respect and rapport, observers would need to note whether student interactions are characterized by conflict, sarcasm, or put-downs (an unsatisfactory rating for the teacher); whether students, in general, refrain from disrespecting one another (a basic rating); whether student interactions are, in general, polite and respectful (a proficient rating); or whether students demonstrate genuine caring for one another and monitor one another's treatment of peers (a distinguished rating).
4. Participants learn how to calibrate their judgments against those of their colleagues. For

example, one observer might interpret interactions in a classroom as representing basic performance, whereas another might see them as proficient. There are many reasons for such differences. One observer might simply have missed something important in the classroom, or the two observers might have slightly different ways of interpreting their evidence. But whatever the reason, it's important they discuss the situation so that they can, in the future, make consistent judgments.

Finding Time for Professional Conversations

A second challenge for administrators is finding time to conduct meaningful observations and engage in professional conversations about practice. However, even in the traditional system, principals need to devote time to the evaluation process—despite the fact that it often produces few benefits. In the words of an educator with whom we've worked, "It doesn't take any longer to do this process well than to do it poorly, so why not do it well?" What better use of a school leader's time than to engage teachers in conversations about practice?

Evaluator-teacher conversations, when conducted around a common understanding of good teaching—and around evidence of that teaching—offer a rich opportunity for professional dialogue and growth. We can't create more hours in the day, but careful setting of priorities and judicious scheduling of both observations and conferences can make the best use of the time available. Moreover, unless a district's negotiated agreement forbids it, brief and informal drop-in observations yield plenty of information for reflective conversation and require far less time than formal observations do.

A Thoughtful Approach

Abundant evidence from both informal observation and formal investigation indicates that a thoughtful approach to teacher evaluation—one that engages teachers in reflection and self-assessment—yields benefits far beyond the important goal of quality assurance. Such an approach provides the vehicle for teacher growth and development by providing opportunities for professional conversation around agreed-on standards of practice.

Endnote

¹ The Framework for Teaching divides the complex activity of teaching into 22 components clustered in four domains of teaching responsibility. Each component defines one aspect of a domain; two to five elements describe a specific feature of a component. To view the framework, visit www.danielsongroup.org/theframeteach.htm.

Charlotte Danielson is an educational consultant based in Princeton, New Jersey. She has taught at all levels and has worked as an administrator, curriculum director, and staff developer. She is the author of *Enhancing Professional Practice: A Framework for Teaching* (ASCD, 1996, 2007) and *Teacher Leadership That Strengthens Professional Practice* (ASCD, 2006); charlotte_danielson@hotmail.com.
