Why Complex Teacher Evaluations Don't Work

Mike Schmoker
Here they come: those complex, bloated, evaluation templates that are now being dumped on teachers and administrators. These are supposed to make schools perform better.

Once again, we are rushing into a premature, ill-conceived innovation—without any solid evidence that it promotes better teaching. These jargon-laced, confusing documents are to be used to evaluate or even to compensate teachers on the basis of multiple, full-period, pre-announced classroom observations. Each observation is to be preceded and followed by meetings between teachers and administrators that will require enormous amounts of time, paperwork, and preparation. Like so many past reforms, this one will be launched nationally, like a bad movie, without being piloted and refined first. (Imagine if we did this with prescription drugs.) It will consume a disproportionate share of precious training time and promote misguided practices that could endure for the next decade. Rather than improve schools, it will only crowd out and postpone our highest, most urgent curricular and instructional priorities.

Don’t misunderstand me: Teacher observation and evaluation are among the strongest components of effective school-improvement efforts. If you visit classrooms across the nation (as many of us do), you know that most teaching is at odds with some of the most obvious elements of sound practice. But these frameworks aren’t the solution. They lack clarity and focus, and their use should be postponed on the basis of their sheer bulk (most are dozens of pages long) and their murky, agenda-driven language.

In February, *The New York Times* reported that one of these frameworks contains an astonishing 116 “subcategories” by which educators’ lessons are to be assessed. I can only imagine teachers, whose morale is already at a record low, encountering these unwieldy instruments and the anxiety they will provoke.

Done right, teacher evaluation could ensure precisely the kind of systematic action that would guarantee immediate improvement, i.e., by clarifying a minimal set of the most essential, widely known criteria for effective curriculum, such as rich content taught largely thought literacy activities and sound instruction.

Once clarified, evaluation would then focus on only one or two elements at a time, with multiple opportunities for teachers to practice and receive feedback from their evaluators. Teachers’ progress and performance on these criteria would be the basis for evaluation.

Jim Collins, the business consultant and author of *Good to Great*, and the organizational-improvement expert Marcus Buckingham discovered that the performance and morale of both employees and managers skyrockets when managers:

- Severely reduce the number of criteria by which they judge an employee’s performance; and
- Have “crystal clarity” for those very few criteria, abandoning any language that could confuse a practitioner.

Teachers need assurances that we will never, ever require them to pore through dozens of bewildering boxes and bullets about how they should perform. Policymakers
have yet to learn that less is more with respect to strategic planning, our (still-gargantuan) standards documents, or our ever-expanding and exotic menus of programs and professional-development offerings. And now teacher-evaluation frameworks.

One popular multi page framework requires that lessons be taught with “simultaneous multi sensory representations” during the lesson and “facilitation…that results in students’ application of interdisciplinary knowledge through the lens of local and global issues.” Another framework—in similarly mangled language—requires that lessons “reflect understanding of prerequisite relationships among topics and concepts and a link to necessary cognitive structures.” I guarantee that is not the kind of advice average teachers need to improve their lessons. Moreover, most of these frameworks insist—against all research and evidence to the contrary—that teachers must provide lessons that include special materials for each individual student or subgroup, all while addressing dozens of other criteria.

We’ll never improve instruction this way. Here’s the alternative.

First, we should do everything in our power to ensure that there is a clear, coherent curriculum in place before
we attach high stakes to any evaluation. The absence of such a curriculum explains a great portion of the aimless, ineffective lessons we see in our schools. In addition, this curriculum must include generous amounts of what is now—finally—being emphasized in the “three shifts” that capture the essence of the English/language arts common core, i.e., daily opportunities to read, discuss, and write. These should all be grounded in evidence found in high-quality, content-rich texts across the disciplines. This simple, timeless emphasis is the key to success on tests, in college, and in careers. It is nowhere to be found, however, in our most popular evaluation templates.

Without such a curriculum, instruction inevitably devolves into the kinds of inane worksheets, group activities, and misguided practices that now predominate in our schools.

Once such a curriculum is in place, we should evaluate teachers on whether they are actually implementing and improving their curriculum in teams, with their same-course colleagues.

“...most of these frameworks insist—against all research and evidence to the contrary—that teachers must provide lessons that include special materials for each individual student or subgroup, all while addressing dozens of other criteria.”

“Done right, teacher evaluation could ensure precisely the kind of systematic action that would guarantee immediate improvement.”

Finally, we should observe and evaluate teachers on the basis of (mostly) short, frequent, unannounced classroom visits, using the same, few, age-old criteria. The noted researcher Robert Marzano, among others, exhorts us to regard these as “routine components” of any and every effective lesson:

- Attention and engagement (i.e., steps are taken to ensure that all students are attentive and on task throughout the lesson);
- A clear, well-defined purpose and objective to the lesson; followed by…
- Multiple short segments of instruction; immediately followed by…
- Opportunities for students to process or practice what was just taught, while the teacher checks and monitors to see how well the class has learned; followed by…
- Adjustments to the lesson and the pace of the lesson to ensure that all students, or as close to that as possible, can succeed on each phase of instruction, until they can achieve the objective of that day’s lesson or group project.

These elements, which guarantee improvement, can actually be found in some of the evaluation frameworks. But they are not written clearly or prominently enough to be seen as indispensable priorities. Instead, they are obscured by the dozens of other specious, confusing evaluation criteria that surround them. To reiterate: The observations that are the basis of an evaluation must occur largely unannounced. We can’t afford to repeat the feckless protocols refuted decades ago—those built around pre-announced visits, followed by lengthy pre- and post-conferences.

Until this changes, as the author and teacher-evaluation expert Kim Marshall and others have made so clear, teacher evaluation will continue to be nothing more than what teachers and administrators have aptly called a dog-and-pony show, with one difference: It will be even more confusing and time-consuming.

It is high time that the reform community grows up and learns that schools won’t improve until we put the brakes on untested, overblown initiatives. These prevent us from focusing on the most effective practices long enough for them to take hold.

Clear, minimalist, priority-driven teacher evaluation could play a central role in ensuring that such practices become the norm. If they do, we will beyond any doubt hasten the improvement of schools in virtually any setting.