Making Teachers
BETTER
Not
BITTER

Balancing Evaluation, Supervision, and Reflection for Professional Growth

TONY FRONTIER | PAUL MIELKE

Hawker Brownlow
EDUCATION

Making Teachers
BETTER, NOT BITTER

Balancing Evaluation, Supervision, and Reflection for Professional Growth

The Need for Balance

"This process isn't making me better. It is making me bitter."

— Cindy Alexander, Middle School Teacher

"We've spent all our time trying to figure out how to make this process more manageable when we should be asking how to make it more meaningful."

— High School Principal

A familiar scenario unfolds in schools each year. It is the final day of state testing. Students meticulously fill in bubbles on answer sheets. They respond to a few open-ended questions. They are not exactly sure how the test will be scored, but they know it is important and that they will be judged based on the results. Later that day, the responses are collected and sent away to be evaluated.

After a few weeks, students forget that they have even taken the tests. Then, near the end of the school year, each student receives a document—a report that judges the student's competence in several domains. Students look at the scores, but they don't know exactly how they were calculated or what they mean. Without this understanding, they (or their parents) reach a broad generalization: *I am good at math*, or *I am OK at science*, or *the test must not have been fair.*

Teachers and administrators understand the realities of student accountability, but they question the benefits of such large-scale standardized student assessments. After all, accountability tests take time away from student learning,

1

the results provide little information that students can use to improve, and the process forces students to passively accept a judgment from an outside source on the basis of a sliver of their performance from the entire school year.

Unfortunately, we believe that in our current push for more rigorous systems of *teacher* accountability, we are in danger of duplicating the flaws of systems of high-stakes *student* accountability. Three days of high-stakes testing does not improve student learning, and three days of high-stakes evaluation does not improve teacher performance. Like high-stakes student assessment, teacher evaluation is seen by many as an occasional event that is disconnected from day-to-day teaching and learning, consumes too much time, produces results that do not help teachers improve their performance, and places those who are supposed to benefit from the system in a passive role as mere recipients of an external judgment. Given the amount of time, effort, and energy invested in systems of evaluation, it is critical that those systems provide benefits that empower and build—rather than merely measure—teacher capacity to support student learning.

This book is not about whether teacher evaluation is *good or bad*. This book is about ensuring that all stakeholders are clear about what teacher evaluation *is and is not*. It is about transcending the misconception that teaching is easy—and that anyone who struggles to address a single student's needs simply lacks the reward or consequence to do so. It is about understanding why evaluation will never improve teacher performance and what to do about it. It is about eliminating the dangerous notion that teachers become expert at their craft after a few years of practice. It is about developing processes of supervision that are less reliant on supervisors. It is about helping stakeholders understand that using comprehensive frameworks of effective teaching to guide personal reflection and changed practice—rather than solely to evaluate—carries the greatest potential to help teachers improve and feel valued. And finally, it is about changing a system that has grown woefully out of balance—with its emphasis on visits, forms, and deadlines—into a system that is productive and supportive of teachers and the students they serve.

The Other Sides of the Pyramid

Our perception of what something is often has more to do with our perspective than reality. This concept is important when considering how comprehensive frameworks for instruction are currently used in schools.

Consider this anecdote. Three nomads converge at a small marketplace, far from where they usually trade, to exchange some goods. Despite the long, harsh journey, the trip is worth it; each nomad has brought goods that cannot be acquired in the others' native region. One has come from the North, one from the West, and the third from the South. As night falls, they meet near the marketplace, start a fire, and talk about their journeys.

Eventually, the conversation turns to the great pyramid that stands, in the distance, between each of their native regions. The nomad from the North says how much he admires the pyramid's beautiful red walls. The nomad from the West laughs and asks if he meant to say yellow. The nomad from the South says they are both wrong; the walls are clearly green.

The nomad from the West claims that according to legend, most of the riches inside the pyramid were placed along the western wall, as a tribute to the good people of the West. The others argue that it was their side of the pyramid that was the most significant. Voices are raised; fingers are pointed.

The nomad from the South claims that not only did his side of the pyramid have the greatest significance, but the wall facing south was the most important because without it, the other sides of the pyramid could not stand. A shouting match ensues.

The argument escalates; no one concedes. Despite desperately needing the goods that each has brought, they are all so angry that they decide they do not want to do business with people whose judgment is so deficient. They break camp in the darkness.

As they head back down their respective trails, the sun rises and the pyramid comes into view. Each nomad looks at his side of the pyramid, satisfied that he was right and the others were wrong. They never do business together again.

Each nomad was correct in articulating his perception of the color of one wall of the pyramid. Unfortunately, each was wrong in defending his perception of what

the other walls *were not.* The truth is the pyramid was built with three types of stone, and the walls were different colors. Each nomad was right about a portion of the pyramid but failed to acknowledge that the other walls could be different.

Furthermore, each was correct in articulating the significance of *his* wall to the people in his region. But this "all or nothing" reasoning failed to acknowledge why the other walls might be significant to others. Unfortunately, each nomad was wrong in his analysis of what part of the structure was *most* important. Although each wall was equally important, the portion that was most important was not a wall at all but the single foundation that supported each wall.

Without a solid foundation, a structure cannot be built. Without balance among the walls, the structure cannot stand.

The pyramid anecdote is analogous to much of the debate we've seen unfold around teacher evaluation. Teachers, administrators, and boards argue about the details of comprehensive frameworks for instruction and the evaluation systems that are often attached to them, yet the purpose of the system is rarely articulated. Furthermore, the debate about why teacher evaluation is so important seems to miss the notion that a solid foundation of trust and credibility is essential if the system is going to stand.

There is one important difference between the pyramid analogy and teacher evaluation. The nomads knew the structure was a pyramid; they were quibbling about the color and the value of *their* wall. We seem to be sitting at the base of the wall that has clearly been labeled as "evaluation," yet we are oblivious to the fact that other walls even exist. What are these other walls that must be present to bring balance to the evaluative component of comprehensive frameworks for instruction?

This era of high-stakes accountability for teachers has spurred an unprecedented emphasis on teacher evaluation. In general, this focus has been on technical, transactional components of the evaluative process, such as the number of administrative visits, which forms need to be completed, and how various scores result in a single rating. We've sat so long at the base of the pyramid—staring at the imposing wall of evaluation—that we've forgotten why we convened at the pyramid in the first place. What started as a conversation about quality and

accountability has devolved into negotiations about visits, tracking forms, and online management systems.

Furthermore, we seem to have forgotten that the other walls exist. By itself, *evaluation* is not a pyramid. It is a two-dimensional triangle balancing on a thin edge that cannot stand. An empowering system of *supervision* and processes that support meaningful *reflection* are equally important walls. We need to clarify the purpose, distinctions, and interrelationships among *evaluation, supervision,* and *reflection* if we are to use comprehensive frameworks in a manner that makes teachers better, rather than bitter.

The Potential of Comprehensive Frameworks for Effective Practice

Much of the discussion of new systems of evaluation has focused on the underlying frameworks that states, districts, or schools have adopted. What follows is a brief overview of some of the more popular teaching frameworks. Each is research-based and uses a standards-based approach to evaluation, with specific criteria described along a continuum of quality. The rubrics can be used for judgment and evaluation, as well as formatively to establish a common language for individual goal setting and professional collaboration. None of the frameworks uses a checklist approach whereby every element needs to be present in every lesson; instead, they honor the systemic and complex nature of teaching and learning.

The Danielson Framework for Teaching

Based on her research for Educational Testing Service on evaluating preservice teachers, Charlotte Danielson (1996, 2007) developed a framework that sought to capture the complexity of teaching in a manner that was practical yet didn't reduce teaching to a set of simple steps. The Danielson Framework divides teaching into four domains: Planning and Preparation, Classroom Environment, Instruction, and Professional Responsibilities. Within these domains are a total of 22 components of effective practice. These components are further divided into a total of 76 elements. The Danielson Framework uses a four-point rating scale to evaluate teacher performance: Distinguished, Proficient, Basic, and Unsatisfactory.