

LEARNING ON THE JOB: Teacher evaluation can foster real growth

Evaluations of practice using research-based standards multiple times throughout the year can provide a focus for professional development, and feedback from evaluations can encourage self-reflection and meaningful conversations focused on classroom practice among educators.



**By Gary W. Ritter and
Joshua H. Barnett**

My very first teacher observation was during my first semester of student teaching in a resource room for kids with learning disabilities. I remember it well because it was traumatic. I hadn't previously met the man who came to do my observation; he was part of the college faculty in the education department but not my instructor.

As the man observed, my lesson spiraled downhill in a sort of out-of-control, catastrophic disaster of epic proportions. At least, that's what it felt like at the time. The kids totally didn't get it, and I wasn't on top of their impulsive and distractible behaviors. When he pulled me into the library so he could talk about my lesson, I knew that it hadn't gone well and as soon as he started talking, I burst into tears.

— Sarah (2011)

Sarah, a kindergarten and 1st-grade teacher, did not enjoy her first evaluation. Very few people enjoy being evaluated. While Sarah's evaluation experience was certainly not unique, there also must be many examples of positive and productive evaluation experiences. Who has had a great evaluation? Who

GARY W. RITTER (garyr@uark.edu) is a professor and endowed chair in education at the University of Arkansas, Fayetteville, Ark. **JOSHUA H. BARNETT** is a senior vice president of research and evaluation for the National Institute for Excellence in Teaching, Santa Monica, Calif.

walked away from their teacher evaluation/review thinking, “Now I’m going to be a better teacher.” While this is an ambitious goal for a teacher evaluation, it doesn’t seem impossible. After all, for the past five years, policy makers in several states have been reshaping teacher evaluation systems in the hopes of improving teacher quality across the board.

In the world of education policy, there is no shortage of ideas about how to fix education — from more resources to fewer resources to more equitable resource distribution. While the yellow-brick road is yet to be discovered, many policy makers, including those in the Obama administration, were drawn to the idea that improved teacher evaluation could play an important role in improving outcomes for students in the U.S.

Teacher evaluation in practice

To learn how meaningful and thorough teacher evaluation systems could actually improve teacher skills and effectiveness, we interviewed about 50 teachers and policy makers engaged in teacher evaluation reform. Specifically, we sought out two sets of educators:

- Educators with experience in a school that has adopted the TAP System, a long-standing, comprehensive school reform model; and
- Educators working in a state that actively engaged in teacher evaluation reform strategies in an effort to win the federal Race to the Top grant competition.

The National Institute for Excellence in Teaching (NIET) introduced TAP: The System for Teacher and Student Advancement in 1999. In addition to TAP, NIET supports states and districts with a range of educator effectiveness strategies. TAP aims to offer career advancement and leadership opportunities for educators, a fair and transparent evaluation process linked to job-embedded professional development, and performance-based compensation. The system is built around four core elements:

- Multiple career paths;
- Ongoing applied professional growth;
- Instructionally focused accountability; and
- Performance-based compensation.

In TAP schools, skilled teachers can serve as master and mentor teachers, receiving additional compensation for providing high levels of support to career teachers and increasing instructional effectiveness across the faculty. Master and mentor teachers form a leadership team, along with administrators, to deliver school-based professional support and conduct

evaluations with a high level of expertise. These master and mentor teachers lead weekly cluster group meetings where all teachers examine student data, engage in collaborative planning, and learn instructional strategies that have been field-tested in their own schools.

Most important for this study, multiple trained observers, including principals and master and mentor teachers, observe TAP teachers in classroom instruction several times a year, using the TAP Teaching Standards rubric, which examines multiple dimensions of instructional effectiveness. Evaluators are trained and certified, and leadership teams monitor the reliability and consistency of evaluations in their schools. These classroom evaluations are complemented by analyses of student achievement growth. Evaluation results are used as formative feedback in one-on-one mentoring sessions and guide planning for cluster group meetings.

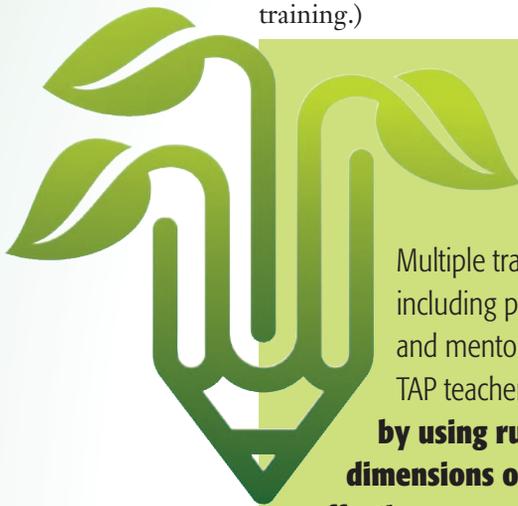
When done well, evaluation is not punitive, it is not an HR function, but it is actually professional development.

Meaningful teacher evaluation underlies nearly every component of the TAP model. Teachers and school leaders involved in TAP are well-situated to provide some perspective on the potential benefits of teacher evaluation reform. A variety of researchers have examined the academic effect of TAP (Algers Charter School Association, 2011; Daley & Kim, 2010; Hudson, 2010; Eckert, 2013; Mann, Leutscher, & Reardon, 2013). However, our focus is to understand more about how this system can change teacher performance.

At the same time, Tennessee teachers and policy makers also have learned lessons related to teacher evaluation that are worth sharing. In July 2011, Tennessee became one of the first states to implement a comprehensive, statewide, student outcomes-based educator evaluation system. This implementation was a key tenet of Tennessee’s First to the Top Act, adopted by the General Assembly as a response to the federal Race to the Top competition. This legislation established the parameters of a new teacher and principal evaluation system. In this system, a substantial portion of the evaluation was to be based on student achievement data, with the remainder to

be determined through qualitative measures, including teacher observations.

Policy makers in Tennessee adopted the TEAM (Tennessee Educator Acceleration Model) rubric as the state model and also voted on a number of key components of implementation, including a minimum number of annual observations. Districts had the option of choosing an alternate evaluation system that met the basic requirements set by the state. In summer 2011, the Tennessee Department of Education worked with NIET to provide a four-day training for all evaluators across the state. NIET trained more than 5,000 evaluators intensively in the state model to ensure they understood the distinction between differing levels of performance. (Districts using alternative instruments delivered their own training.)



Multiple trained observers, including principals and master and mentor teachers, observe TAP teachers several times a year **by using rubrics for several dimensions of instructional effectiveness.**

Tennessee's teacher evaluation system has undergone a great deal of change in recent years. Thus, the perspective of those in the trenches in Tennessee will be quite valuable to the question addressed here as many have directly experienced earlier evaluation systems as well as the improved version.

What teachers said about evaluation

The overarching question is whether the exercise of evaluating teachers can actually make teachers better. Four themes emerged from our investigation of the potential benefits of enhanced and comprehensive teacher evaluation strategies.

Finding #1: Business-as-usual fails. We didn't expect to hear so much about the former system in our conversations with educators. However, each conversation was filled with multiple statements indicating how things used to be, and they were not the good ol' days. Most educators clearly described their former evaluation system as not useful.

An elementary school teacher at a TAP school in

Louisiana said that, before TAP, "evaluation was just something that we had to do. The principal would stay for a dog and pony show lesson. . . . We didn't get feedback, and I didn't grow."

A Tennessee principal described her previous teacher evaluations as proforma exercises with an evaluation instrument that wasn't at all rigorous. The conversations following the observation were brief, something along the lines of: "Here are your scores. Can you sign here?" Thus, there was little differentiation between teachers, and the process did not encourage any sort of professional accountability.

While this type of evaluation seems meaningless, we would have hoped at least that it was innocuous. However, some respondents suggested these evaluations could actually have negative repercussions for teachers across the effectiveness spectrum. According to a principal from a Texas school, "Stellar teachers would resent the traditional evaluation systems because they would simply receive high marks without any feedback to help them boost their craft." Moreover, these highly effective teachers realized that the high marks they received were not truly meritorious since all the teachers in the school were receiving high marks as well.

On the other side, teachers who were not so confident in their ability would resent the evaluations as a gotcha strategy. This negative feeling would occur even though almost no teachers ever felt the sting of an evaluation since virtually all teachers would receive satisfactory ratings. Indeed, in some ways, traditional evaluation systems incorporated the worst of all worlds: There was no true professional accountability or support even as struggling teachers viewed the systems as being quite punitive.

Finding #2: Evaluation can provide a focus for professional development. While many are optimistic about the potential for effective professional development, there is little convincing evidence to guide school leaders choosing professional development strategies. Aligning professional development to the evaluation system is one promising strategy for creating effective development for educators. Educators operating in environments with meaningful evaluation systems say professional development can be most helpful when focused on all aspects of the evaluation system — from the observation rubrics to the data gathering to the feedback.

One of the four key components of TAP is ongoing applied professional development. Several educators said TAP professional development can be effective precisely because it uses an evidence-based educator effectiveness rubric, multiple classroom observations, and analysis of student data. According to a 6th-grade,

lead teacher in a TAP school, this focused professional development is far better than what she had experienced in other schools. Her assessment of traditional professional development is not good, she said:

Prior to TAP, I gave an open invitation meeting for teachers to come in and give talks on things they were experts on . . . I was just hoping that we would interact since there was no professional development ongoing . . . In the best case scenario, we would take home one or two nuggets we could put in our teacher toolkit.

Not surprisingly, she spoke much more positively about the embedded professional development in TAP schools. “Everything we do now is customized to our needs based on data and evaluation,” she said. The theory of action for TAP is that ongoing applied professional development will be created and delivered through local teacher leaders who serve in TAP schools as master and mentor teachers. According to a former teacher and principal, school leaders who create meaningful evaluation systems are changing the paradigm of evaluation. “When done well, evaluation is not punitive, it is not an HR function, but it is actually professional development,” she said.

Finding #3: Meaningful evaluation creates a space for meaningful feedback. Meaningful teacher evaluation can create opportunities in schools for teachers to engage in meaningful conversations focused on classroom instruction and student achievement. These conversations can encourage teachers to be more self-reflective and seriously examine data on student learning. According to most of those interviewed, the feedback resulting from observation postconferences may be the most important contribution of improved teacher evaluation systems, and teachers benefit from the opportunity to reflect. The result is a school setting in which teachers and leaders are having regular conversations about improving instructional practice and student learning.

In schools with meaningful evaluation, according to our respondents, the evaluation process is comprehensive and goes well beyond the classroom observation. The observation is followed by a postconference with feedback, a follow-up in the near future, and, ideally, self-evaluation by the teacher. Teachers in schools with meaningful evaluation systems appreciate the postconference and feedback because it gives them strategies to improve their craft (and their evaluation scores) in the future. A former teacher and school leader who now works with TAP schools described the feedback this way: “We will analyze together and reflect on one thing that the teacher did really well and one thing that the teacher should reflect upon. Every teacher will have an area of reinforcement and refinement.”

In some ways, traditional evaluation systems incorporated the worst of all worlds: **There was no true professional accountability or support even as struggling teachers viewed the systems as being quite punitive.**

One education policy maker involved in revamping the teacher evaluation system in her state said, “All people need feedback to improve practice; in teaching, we have not done a great job of providing feedback. Improved teacher evaluation is less about the scores than the process. In a system that works, the agent of change will be the feedback associated with the score.” One school leader said the two most important drivers of improved instruction in his school were the interaction among teachers (actively visiting each other’s classrooms and observing) and the cluster meetings focused on student results and instructional strategies. These interactions and meetings provide a setting where these productive conversations can occur.

Finding #4: Teachers revealed an appetite for evaluation. We did not expect teachers to respond favorably to being evaluated. However, educators said their peers were willing to accept rigorous evaluation as long as the evaluation was productive and the environment was positive. When the observations were frequent and the rubrics thoroughly explained in cluster meetings, teachers were more willing to view the ratings as meaningful. If an evaluator comes in once per year, the observer is simply trying to record a rating or a score; multiple visits and feedback suggest a genuine interest in helping the teacher get better.

Additionally, rigorous evaluations are far more likely to be welcomed in an environment of trust and shared responsibility. Educators said there are numerous ways to foster an environment of trust. In TAP, for example, teachers trust that they will receive meaningful feedback and opportunities for improvement after each observation. Moreover, teachers understand and trust the usefulness of the evaluation rubric, and they are confident that the master teachers and evaluators fully understand the rubric and its connection to improved instruction. Classroom teachers respect the master teachers because it is clear that these master teachers are working incredibly hard, they are themselves effective educators, and they are subject to evaluation by the very same rubric.

Join the conversation

facebook.com/pdkintl
[@pdkintl](https://twitter.com/pdkintl)



Teachers become more open to evaluation as they see professional growth as a result of the evaluation system.

Finally, teachers become more open to evaluation as they see professional growth as a result of the evaluation system. When these conditions are met, teachers are willing to be scrutinized by a system that was not built in Lake Wobegon, where everyone is above average!

Conclusion

In 2010, many education policy observers believed that improved teacher evaluation with a real focus on student outcomes (even the dreaded student test scores) and with the support of generous financial incentives would be a reform with some staying power. Fast-forward five years and the strong backlash against testing has dominated the teacher evaluation

discussion. However, advocates for improved teacher evaluation might view recent developments as a policy opportunity — policy makers may well be receptive to a teacher evaluation strategy focused on observation and feedback in addition to student achievement data like the ones described in this article.

In this study, we interviewed nearly 50 teachers, administrators, and policy makers in order to elicit and understand lessons to be learned from those schools and states that have approached teacher evaluation reform proactively. We conclude that serious implementation of professional, rigorous, and comprehensive teacher evaluation systems represents a promising school improvement strategy.

Education policy makers around the nation, particularly those considering teacher evaluation reforms in their own states, may benefit from the lessons derived from the educators in our sample. In those settings, where new forms of teacher evaluation are being tested, let's make sure that the school leaders get together with researchers to rigorously assess how these programs affect student learning. **K**

References

- Algiers Charter School Association. (2011). *Annual report 2011*. New Orleans, LA: Author. www.tapsystem.niet.org/publications/algiers_charter_schools_association_annual_report_2011.pdf
- Daley, G. & Kim, L. (2010). *A teacher evaluation system that works* (NIET working paper). Santa Monica, CA: National Institute for Excellence in Teaching. www.tapsystem.niet.org/publications/wp_eval.pdf
- Eckert, J. (2013). *Increasing educator effectiveness: Lessons learned from Teacher Incentive Fund sites*. Santa Monica, CA: National Institute for Excellence in Teaching. www.niet.org/assets/increasing-educator-effectiveness-lessons-learned-from-teacher-incentive-fund-sites.pdf
- Hudson, S. (2010). *The effects of performance-based teacher pay on student achievement*. Discussion paper for the Stanford Institute for Economic Policy Research. Stanford, CA: Stanford University. <http://bit.ly/1n001TG>
- Mann, D., Leutscher, T., & Reardon, R.M. (2013). *The system for teacher and student advancement: An evaluation of achievement and engagement in Louisiana*. Santa Monica, CA: National Institute for Excellence in Teaching. www.niet.org/assets/PDFs/interactive-louisiana-student-achievement.pdf
- Sarah. (2011). *A horror story from the trenches of student teaching*. *Confessions of an untenured teacher*. [Web log]. <http://untenured.blogspot.com/2011/02/horror-story-from-trenches-of-student.html>



“Think inside the box, Roger.”