

**T**eacher evaluation systems should be linked to professional development if they are to provide the crucial feedback teachers need to analyze their work and receive targeted support. That's one of the goals of TAP: The System for Teacher and Student Advancement. Managed and supported by the National Institute for Excellence in Teaching (NIET), TAP was introduced in 1999 to attract, motivate, and retain talent in teaching by providing opportunities for career advancement, professional development, evaluation and performance-based compensation. Each of TAP's four elements reinforces the others to form an interconnected support system for educators with the ultimate goal of improving learning gains for all students.

A powerful dimension of TAP's support structure is the cadre of teacher leaders TAP develops in each school, made up of master and mentor teachers. Master and mentor teachers, along with the principal, drive instruction in the school by analyzing student data, creating achievement plans, leading professional development and evaluating and supporting other teachers. Because active collaboration is essential, teachers must vote to implement the reform on their campus. TAP is currently at work in some 380 schools across the U.S., involving more than 20,000 teachers and 200,000 students.

TAP's approach to teacher evaluation has two equally important goals: accurately measuring teacher performance and improving teacher skills through individualized, intensive support.

Those goals are connected by the research-based TAP Teaching Skills, Knowledge, and Responsibilities Performance Standards. These rubrics are broad and detailed enough to capture essential elements of teaching and to increase improvement. Teachers are scored using a 1-5 scale on indicators of effective instructional practice crossing three domains: instruction, designing and planning instruction, and the learning environment. A score of 1 is unsatisfactory, 3 is proficient, and a 5 is exemplary. At the end of each year,

**JASON CULBERTSON** is chief learning officer and executive vice president of the National Institute for Excellence in Teaching, Santa Monica, Calif.

teachers also receive scores on additional indicators in a fourth domain — responsibilities — that recognizes teachers for their efforts to improve teaching.

TAP’s evaluation system includes multiple classroom observations each year by multiple trained and certified evaluators, including principals or other administrators, master teachers, and mentor teachers. Announced observations are preceded and followed by in-depth teacher conferences during which the evaluator and teacher examine the lesson to identify a strength (area of reinforcement), a weakness (area of refinement), and a specific plan for improvement. Since evaluators know the teachers and coach them on a regular basis, they have a more robust context for selecting areas for reinforcement and refinement after observations. Furthermore, they have more opportunities to provide teachers with intensive follow-up support in those areas after the observations.

One of the most important results from this process is that educators build a common language around what effective teaching looks like and recalibrate their expectations to create significant room for growth for even the most accomplished teachers. By defining proficiency as a 3 on a 1 to 5 scale, nearly every teacher is provided with stretch goals and participates in a process of continuous improvement.

In addition to scores on a teacher’s classroom practice — segmented as skills, knowledge, and responsibilities, or SKR — each teacher receives a value-added score that measures the teacher’s effect

on student learning growth. When teachers demonstrate strong instructional skills as measured by the TAP Teaching Standards, their students show higher academic growth regardless of previous achievement and socioeconomic status. (National Institute for Excellence in Teaching, 2012, p. 12).

The TAP Teaching Standards and methods for observation represent a major departure from the “meets/does not meet” approach that characterizes most evaluation systems. A 2011 report by the Center for American Progress found that “in most places, teacher evaluations are infrequent; are based on scant evidence; rely on crude instruments; include few reliable quality controls; fail to use adequately trained evaluators; provide almost no useful feedback to teachers; and yield vastly inflated performance ratings” (Jerald & Van Hook, 2011, p. 7).

### **‘A form, a checklist’**

Master Teacher Sara Greene of Holston Middle School in Knoxville, Tenn., remembers her evaluations before TAP. “It was a form, more or less a checklist,” said Greene, a science teacher. “They checked to see if your room looked good and if your students responded well to you. I think there was an item to check for dressing appropriately. But there was nothing there that would help me be a better teacher.”

In contrast, the TAP Teaching Standards contain sufficient detail for teachers to acquire a clear under-

**FIG. 1.**  
**Levels of effectiveness for standards and objectives**

#### **Exemplary (5)**

- All learning objectives and state content standards are explicitly communicated.
- Subobjectives are aligned and logically sequenced to the lesson’s major objective.
- Learning objectives are: (a) consistently connected to what students have previously learned, (b) known from life experiences, and (c) integrated with other disciplines.
- Expectations for student performance are clear, demanding, and high.
- There is evidence that most students demonstrate mastery of the objective.

#### **Proficient (3)**

- Most learning objectives and state content standards are communicated.
- Subobjectives are mostly aligned to the lesson’s major objective.
- Learning objectives are connected to what students have previously learned.
- Expectations for student performance are clear.
- Most students demonstrate mastery of the objective.

#### **Unsatisfactory (1)**

- Few learning objectives and state content standards are communicated.
- Subobjectives are inconsistently aligned to the lesson’s major objective.
- Learning objectives are rarely connected to what students have previously learned.
- Expectations for student performance are vague.
- Few students demonstrate mastery of the objective.

standing of what performance looks like at various levels of expertise in a range of classroom practices and skills. For example, the standards and objectives indicator lists five attributes that must be present consistently to reach exemplary performance, including communicating learning objectives and state standards to students; the connection of the learning objectives to a student's own knowledge; and the expectations for student performance. (See Figure 1 for the full list.)

Greene said that before TAP, "I had a good rapport with kids and other teachers. But whatever I did, I didn't have strong student measures, so, honestly, how did I know or have any way to measure how effective I was?"

**The TAP Teaching Standards and methods for observation represent a major departure from the "meets/does not meet" approach that characterizes most evaluation systems.**

Intensive training for TAP starts before staff members adopt the system and continues throughout the year. Sessions are led by master and mentor teachers who explain the instructional practices measured by the TAP Teaching Standards. They model such practices in professional development meetings (known as "clusters") and in the classroom. This is where the core collaboration happens. Teachers analyze student data and learn new instructional strategies to improve student learning. Master teachers select the strategies based on detailed analyses of student achievement data. They are only introduced to teachers in the cluster group after master teachers successfully "field-test" and refine the strategies in classrooms so they can demonstrate student learning gains. After master teachers introduce a new strategy, teachers use it in their own classrooms, then they have cluster meetings appointed with formative assessment data from before and after the strategies were adopted. They discuss how well the strategy worked, then refine it if necessary.

To supplement the cluster group meetings, master and mentor teachers regularly visit teachers' classrooms to provide highly intensive and personalized coaching, ranging from teaching demonstration lessons and modeling specific instructional strategies or skills to team teaching. Mentor or master teachers meet with individual teachers outside of the classroom as well. Validity and reliability of the evaluation system are maintained by ensuring that evaluators receive intensive training followed by certification and annual recertification tests. Evaluators are also

required to undergo training on protocols and data management systems to maintain rating consistency and quality control.

While a significant number of teachers have opportunities to participate in peer observations, experts suggest that their work often is hampered by the lack of a shared understanding of what effective instruction should look like across classrooms. In contrast, the detailed descriptions of practice at various levels of effectiveness in the TAP Teaching Standards enable teachers to build a rich, meaningful, shared language about effective instruction. This process allows them to better collaborate during weekly professional learning opportunities.

Science teacher Greene always furnished lesson plans, but they weren't always tied to the state standards. Through TAP's standards and objectives indicator, she learned how to match the two expectations.

"I never broke down a lesson into what I needed to teach according to what the state standards require, the community needs, and what I want my students to leave the classroom with," Greene said. "Through analyzing student work, I was able to receive feedback on what benchmarks students reached."

The TAP Teaching Standards are also designed to help teachers develop higher-level thinking — an area of need Holston identified. For example, in a lesson on the functions of a cell, a higher-level thinking question would be: Which three organelles do you feel are most important? In this situation, the student would not only have to know what an organelle is, but also would have to be able to compare one to another. Going a step further, a teacher could ask what would happen to the cells if they didn't have those three organelles.

**Reviewing practices**

Just as the TAP Teaching Standards help keep teachers on track toward continuous improvement, they also facilitate an in-depth review of practices across a school. After the 2009-10 year, Holston teachers learned that following two solid years of achieving the highest value-added growth score, they received a lower growth score. Holston educators responded by employing the TAP Teaching Standards' evaluation process, and its professional development facility, to internalize higher standards and develop the capacity to critically reflect on their own performance in key areas. The rubrics also provided a powerful source of ideas for re-evaluating practice to address student needs that may have changed over time.

Through a dynamic exchange of ideas, teachers decided that their overall school strategies in literacy and math were too complicated. By studying assessments, they were able to pinpoint that students had

struggled to understand questions during reading. In math, students required help organizing various mathematical processes to solve problems. As a result, the teachers created a simpler literacy strategy that encouraged teachers to preview what students needed to know at the beginning of a lesson and review the main points afterward. This practice allowed students to better digest and comprehend the information. For math, teachers worked with students to reinforce the mathematical processes for each problem and helped streamline their thought process.

Greene notes that students became so familiar with the strategies that one student asked if he could use both literacy and math strategies during testing time. This was also testament to the strategies' effective transfer to the classroom: Students had achieved mastery.

After this major year of reflection (2010-11) and retooling, teachers learned that their value-added scores soared to a 5 and remained a 5 for the 2011-12 school year. Many schools that go through similar experiences of reassessment resort to outside interventions for the answers. One of TAP's strengths is that teachers and leaders in a building develop responses to questions about low performance and

produce concrete plans for improvement. Every strategy relates to the teachers' own students, allowing teachers to examine their progress along a path and customize strategies for individual students.

The SKR scoring scale of 1-5 also helps teachers

**One of TAP's strengths is that teachers and leaders in a building develop responses to questions about low performance and produce concrete plans for improvement.**

understand progress along a spectrum with plenty of room to grow. Because of the inflated ratings delivered in typical evaluation systems, teachers often expect that they will attain the highest score at the outset. TAP captures teacher performance levels more realistically along a continuum, meaning that a score of 3 represents solid, proficient teaching. This range also allows for feedback to more teachers along a much wider range of performance levels, including experienced and relatively expert teachers who would naturally score at the highest level in systems with fewer categories and consequently receive no useful feedback.

# PDK MEMBERSHIP



CONNECTIONS



RESOURCES



INSPIRATION



We've got what you need.

[www.pdkintl.org/join](http://www.pdkintl.org/join)

Laura Roussel, a former master teacher at Lowery Intermediate School in Ascension Parish, La., who now works as an executive TAP master teacher at the state level, invested time to help teachers fully understand what performance looks like at each level. “We really spent a lot of time in our cluster meetings that first year describing, modeling, and promoting level 3 as solid instruction,” she recalls.

Scoring a 5 in a specific area or overall connotes exceptional or “exemplary” performance. Because “exemplary” is difficult to attain, every teacher is encouraged to continue to improve. This makes the evaluation system relevant, meaningful, and useful to a much larger proportion of teachers in any given school or district.

### Gaining teacher acceptance

Sue Way, a 35-year teacher, principal, and former executive master teacher with the Louisiana Department of Education, recalls that when she first implemented the TAP system as a school principal, her biggest worry was that teachers would not accept the new evaluation system.

“After we spent the first 12 or 14 weeks studying the TAP Teaching Standards, I told teachers to pick any lesson they’ve taught and do the self-evaluation as an informal exercise,” she said. “I just wanted to see where we were. And they were rating themselves all

4’s and 5’s! So, we had to have a talk. I had them compare their self-evaluation ratings with our school’s student proficiency results and asked, ‘If we’re all so wonderful in teaching, why aren’t we off the charts in learning?’ After that, and with more practice, their scores became better calibrated.”

Having multiple trained and certified evaluators helps guard against score inflation and ensures that scores are consistent across the board. During the school year, leadership teams devote at least one meeting per month to discussing issues related to evaluation and analyzing data in order to identify potential ratings consistency problems.

NIET has developed a system to manage, store, and analyze teacher evaluation results, called the CODE system. CODE produces a range of analyses that make it easy for evaluators to spot inconsistencies in scoring. If inconsistencies are found in a rubric indicator, the team has an in-depth discussion, referencing real examples of the indicator from lessons. The result is a shared operational understanding of what constitutes varying levels of effectiveness within the given indicator.

Leadership teams can employ additional strategies to ensure quality evaluations. These can include conducting teamed evaluations or inviting highly experienced evaluators from outside the school, such as executive master teachers or TAP national staff members, to assist in calibrating evaluation scores.

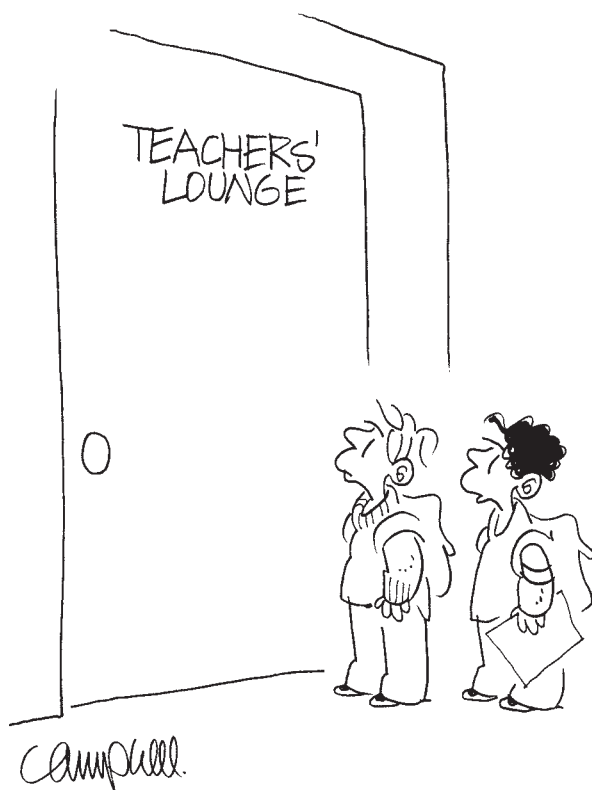
In addition to strong support provided in the school buildings, NIET created the TAP System Training Portal, complete with training modules, strategies, and video lessons scored by national raters, so teachers can gain a deeper understanding of the practices while also learning from TAP educators across the country. The portal also allows teachers to troubleshoot issues and compare their lessons to a national standard, further verifying the validity of their own methods.

TAP shows that an evaluation system will be supported by teachers if it is based on a research-based framework of best practices and built to accurately measure performance and improve teacher skills.

Said Greene, “I cannot imagine finishing my career without knowing what I know about teacher practices, all because of the TAP evaluation process.”

### References

- Jerald, C. & Van Hook, K. (2011). *More than measurement: the TAP system’s lessons learned for designing better teacher evaluation systems*. Washington, DC: Center for American Progress, p. 7
- National Institute for Excellence in Teaching. (2012). *The effectiveness of TAP: research summary 2012*. Santa Monica, CA: National Institute for Excellence in Teaching. p. 12.



“I think it’s some kind of old-fashioned social-networking tool.”