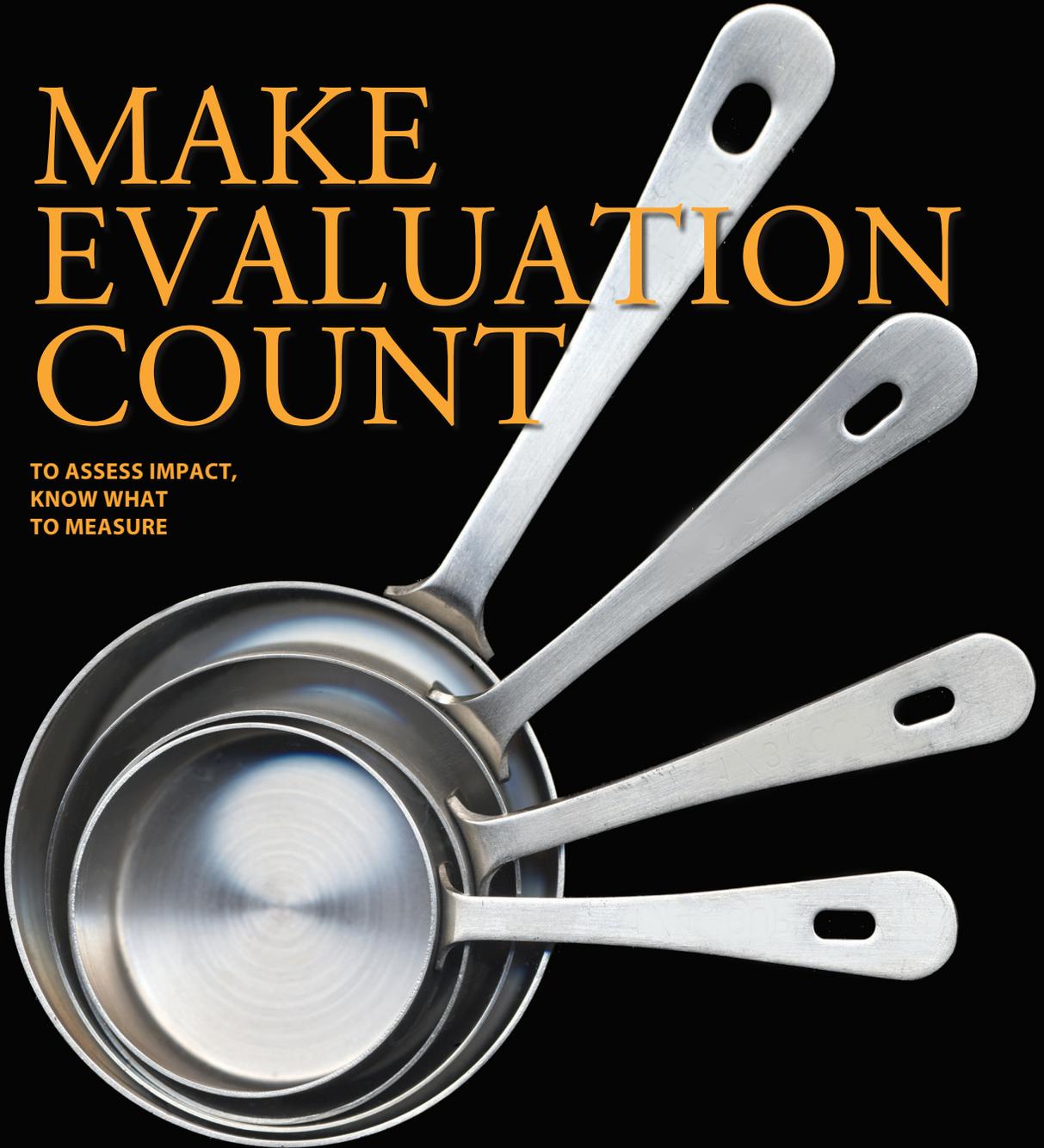


MAKE EVALUATION COUNT

TO ASSESS IMPACT,
KNOW WHAT
TO MEASURE



By Amy Pendray and Jennifer Crockett

In complex educational systems, stakeholders with varied interests often put the greatest value on singular, summative outcomes tied to high-stakes tests. While those summative outcomes are useful and important, an investment in teacher learning intended to improve student achievement is also a treasure worth pursuing.

So how do we, at a systems level, know that our investment in teacher learning is making a difference? How do we ensure that professional learning is impacting teacher practice in a way that leads to improved student outcomes?

As professional development providers for myPD (an online, personalized professional growth system) in Long Beach Unified, a large urban school district in California, we feel a responsibility to wrestle with these questions. We want to deliver high-quality professional learning that ultimately increases student learning. To ensure that all teachers and students benefit from the most effective professional learning we could provide, we have to reflect on our own practices to ensure that we, too, are making a positive impact.

For us, this realization became more pronounced in the context of a broader learning community that extended beyond the borders of our district — Learning Forward’s Redesign PD Community of Practice. With outside eyes looking in on the work to challenge our assumptions and help us deepen our perspectives, we partnered with 20 other districts from across the nation and committed to a problem of practice focused on measuring the impact of our professional development.

As we have grown in our understanding of this work and developed tools to better measure this impact, our partner districts in the community have provided critical and constructive feedback to refine our work.

PREPARING FOR THE WORK

The process began at Learning Forward’s Annual Conference in December 2015, where we learned about Thomas R. Guskey’s *Evaluating Professional Development* (Guskey, 2000). With guidance and support from Learning Forward and McKinsey & Company facilitators, we embarked on a very messy journey in which we began to identify gaps in the way we assessed the impact of professional development on teacher beliefs, knowledge, and skills and how these affect student learning outcomes.

We realized we did not have a way to think through and close the gaps we identified. We knew it was our responsibility, in service to students and teachers, to evaluate the efficacy of our professional learning. Using Guskey’s Critical Levels of Evaluation (see box at right), we analyzed our professional development offerings.

Important trends surfaced. Our measurement of participants’ reactions (Level 1) was very strong. However, we measured use of new knowledge and skills (Level 4) less frequently, and we found challenges on several other levels, indicating a design-implementation gap.

This gap between our intentions in designing and delivering high-quality professional learning and its impact on teacher practice and student learning challenged us to consider adjustments to our approach. We not only needed to evaluate teacher learning, but also follow up with teachers to see how they were using their new knowledge and skills, and, ultimately, determine how the professional

GUSKEY’S CRITICAL LEVELS OF EVALUATION

Level 1: Participants’ reactions.

Level 2: Participants’ learning.

Level 3: Organization support and change.

Level 4: Use of new knowledge and skills.

Level 5: Student learning outcomes.

Source: Guskey, 2002.

learning impacted students.

To deepen our understanding of the work, we filled the next six months with discussions, academic readings, and prototyping and testing new approaches to delivering and evaluating professional development.

Seizing the opportunity for more robust conversations about evaluation, Pamela Seki, assistant superintendent in the Office of Curriculum, Instruction and Professional Development, engaged the entire department in the same reflection process.

Overall, the results were similar to ours, identifying our potential to increase the impact of professional learning and providing the context for an *Evaluating Professional Development* book study to build the common foundation and framework needed to evaluate the department's professional learning. The book study led us to develop a protocol tool to help us move from Guskey's theoretical framework to the practical application in our context.

The tool would help us understand if and when we were intentionally assessing, measuring, and evaluating our professional learning. We wanted to see how all of our professional learning efforts worked together within initiative goals and what additions or adjustments might be required within each professional learning offering to address the appropriate level of evaluation.

THE BUMPY ROAD

Although it was a little bumpy along the way, we realized two things: We needed a formal way to capture the complex thinking we were doing, and we needed to leverage that information to plan comprehensive professional learning that could be evaluated at multiple levels for its efficacy.

Thus we created a prototype of a protocol and evaluation profile matrix to help us determine what to measure at different points within a professional learning program as well as a single professional development offering. After planning the professional learning, we can use the protocol and matrix to reflect on and develop next steps in a professional learning initiative.

We tested the prototype with multiple audiences to get critical feedback and refine the protocol. One particular audience was the beginning teacher support and assessment induction team, which hosts multiple learning opportunities throughout the year.

"The protocol gave me an outside perspective of what our team was doing," said induction support provider Ashley Rhodes, "and made us think about more quantitative evaluation data rather than just going by a feeling that what we were doing was working. It gave us specific measures to consider."

These conversations surfaced the innately subjective way that we had been evaluating the efficacy of our professional learning and challenged us to consider intentional, well-thought-out, and objective measures of our efficacy in supporting teacher learning and student achievement. In some cases, it prompted us to consider building these measures in the profes-

sional learning planning in addition to adapting and revising existing professional development.

Once we refined our work, we tested it with a wider audience. We understood the potential of the process because we had built it, but we wondered if others would find as much value as we had in this reflection. We asked for feedback from a variety of sources within the Office of Curriculum, Instruction and Professional Development.

"When you [Amy and Jennifer] asked me certain questions, it made me reflect on things I had not previously considered evaluating during professional development offerings. It pushed me past the boundaries of what I thought was successful," said Stacy Casanave, English language arts curriculum coach and induction coordinator.

The feedback made clear two critical distinctions that professional learning planners need to make between the types of activities at the heart of professional learning offerings. Instructional activities are best used to help participants understand professional learning content, while evaluation activities are specifically planned methods and processes to gather data to determine if the professional learning is reaching its intended goals.

The importance of clarifying and distinguishing the purpose of each activity is crucial because it is easy for the lines to get blurred. Differentiating between instructional activities and evaluation activities ensures that professional learning planners are on the right track and assessing the pertinent information to determine if program goals are being met.

Not every instructional strategy is used to evaluate professional learning's effectiveness. Some simply move the instruction forward and assist teachers in learning the content. Professional learning planners need to be cognizant of which activities determine the efficacy of the professional development offered — a process that our protocol clarified for us.

MOVING FORWARD

As we continued to use the protocol, small insights along the way led to further refinements in our efforts to better measure all levels of our impact. Because Guskey's critical levels build on one another successively, each iteration of our process and opportunity to reflect on our work gave us a clearer picture of gaps in our professional learning offerings and equipped us with the language and understanding to fill those gaps intentionally and thoughtfully.

Having a defined process that clarified what evidence to collect and how to use it removed the subjectivity upon which professional learning planners rely to make decisions about the effectiveness of their offerings and replaced it with actionable data.

Our collective inquiry around measuring the impact of our professional learning led us to some valuable conclusions. For instance, we learned that evaluating participants' reactions (Level 1) is more than just making sure the participants were happy and had a good time during the professional learning experience.

The context (the physical space of the offering and current mental space of the participants) and process (how the professional learning is structured) specifically affect participants' overall reaction. The protocol helped us uncover the fact that we mostly evaluated for content and did not focus on gathering data on either context or process. Context and process are easily overlooked, yet play a critical role in how participants perceive the quality of professional learning.

Neglecting to assess, measure, and evaluate all aspects of participants' reactions can hinder present and future implementations of learning. In response, we developed and distributed a survey with questions that focused on context and process. The information we gathered helped us redesign the professional learning to meet our participants' identified needs, while building evaluation activities into the day helped us determine the degree to which we were meeting those needs.

Embedded in the same survey were questions that measured organization support and change (Level 3), something we had never even considered assessing. Guskey states, "Information at this level helps us document the organizational conditions [and culture] that accompany success or describe those that might explain the lack of significant improvement" (Guskey, 2000, p. 150).

We identified a clear misalignment between the systems-level professional learning and messaging around our work and site-based implementation efforts. Though we were unable to change course in the midst of the initial professional learning that yielded this data, it has shaped our strategy for partnering with site leaders to ensure coherence and site support for future implementation.

Sparked by these realizations, we were determined to address our challenges in use of new knowledge and skills (Level 4). The readings from Guskey taught us that we needed to allow sufficient time to pass between professional learning and observations of practice to evaluate participants on their use of new knowledge and skills. To accomplish this, we piloted our evaluation of Level 4 on a group of users that had already engaged in professional learning and had been using the new practices in their respective roles for awhile.

We developed a questionnaire that gathered data and information on how participants used their new knowledge over the previous three to nine months. The questions included where the participants felt confident in their use of what they had learned and where they were still feeling challenged by particular skills needed to put what they learned into practice.

The evaluation results provided specific information that allowed us to design follow-up professional learning targeted to participants' needs. We also used the data to update the professional learning content for new participants. Measuring and assessing participants' use of new knowledge and skills (Level 4) was eye-opening, and we will continue this process for each professional development offering we plan.

A COHESIVE AND SYSTEMATIC APPROACH

Professional learning programs have overarching goals that address both student and teacher outcomes. Large initiatives often require multiple professional development offerings in order to reach those goals. The protocol we created enabled planners to look at each offering individually to evaluate the data gathered.

However, a more compelling realization was that using the protocol provided evaluation data that could also be used as a leading indicator (formative) or lagging indicator (summative) for the initiative itself. Like any good road map, leading and lagging indicators allow professional learning planners to make adjustments to their initiative along the way.

As the myPD team tested digital tools to evaluate professional learning, we identified leading indicators, such as quiz results, as well as lagging indicators, including how many sites participated and who facilitated the professional learning. These leading and lagging indicators allowed the team to make course corrections to the overall initiative, thus ensuring a cohesive and systematic approach to planning, implementing, and assessing professional learning's impact. The protocol was a useful tool for planning and reflecting on initiatives and individual offerings both individually and collectively within the larger scope of the initiative goals.

"The power of the protocol that Amy and Jennifer developed is that it moves us from theory to practice," said Nader Twal, program administrator at Long Beach. "It takes something that we all admire — Dr. Guskey's rich work on evaluating the efficacy of professional development — and it gives us a process to calibrate our work around all five levels that he describes.

"It helps us to be intentional and focused in ensuring that not only are we measuring teacher reaction and student outcomes but that we also recognize the important and intermediary measurements of teacher learning, system support, and teacher practice.

"It's iterative and honors the fact that even adult learning can be messy. But much like art, a masterpiece will emerge from the mess. It's about time that we measure what we treasure."

REFERENCES

Guskey, T.R. (2000). *Evaluating professional development*. Thousand Oaks, CA: Corwin.

Guskey, T.R. (2002). Does it make a difference? Evaluating professional development. *Educational Leadership*, 59(6), 45-51.

•
Amy Pendray (apendray@lbschools.net) and Jennifer Crockett (jcrockett@lbschools.net) are program specialists at myPD. ■