Measuring Teacher Professional Learning: Why It's Hard and What We Can Do About It

Case Studies

rppl

In October 2023, the Research Partnership for Professional Learning (RPPL) shared a white paper about measuring teacher professional learning. It was collectively produced by a working group across RPPL's network, centering the voices and practices of organizations working alongside districts, schools, and teachers, understanding that they are well-situated to surface the challenges and needs of practitioners. In the report, we describe the measurement challenges that PL organizations confront and offer views from inside these organizations about where there are opportunities for improvement.

This paired case study set is a collection of real-world examples from each organization that shows how they use and develop innovative, practical, equity-focused measures to address those challenges.

Table of Contents

Teaching Lab

How to Build Comprehensive Frameworks for Data Collection
By Shaye Worthman

Teaching Lab

PL organizations require overarching evaluation frameworks to guide their measurement cycles and determine the impact of their services on teachers and students. The Guskey Framework (2016)[1] is a useful starting point for PL organizations to begin to develop their own evaluation plan that aligns with their theory of action. While the framework is simple and linear, disentangling the correlational and causal relationships between the different levels and the impact they have on one another is not.

The first four aspects of Guskey (participant perceptions, knowledge, mindsets, and enabling conditions) can be reasonably discerned from surveys and assessments, effectively evaluating the most important outcomes, teacher practice and student learning. However, what is increasing in complexity and resource-intensive are classroom observations and student data, including student surveys, student work samples, and formative and summative assessment data as they require time and important human resources to collect, analyze, and interpret.
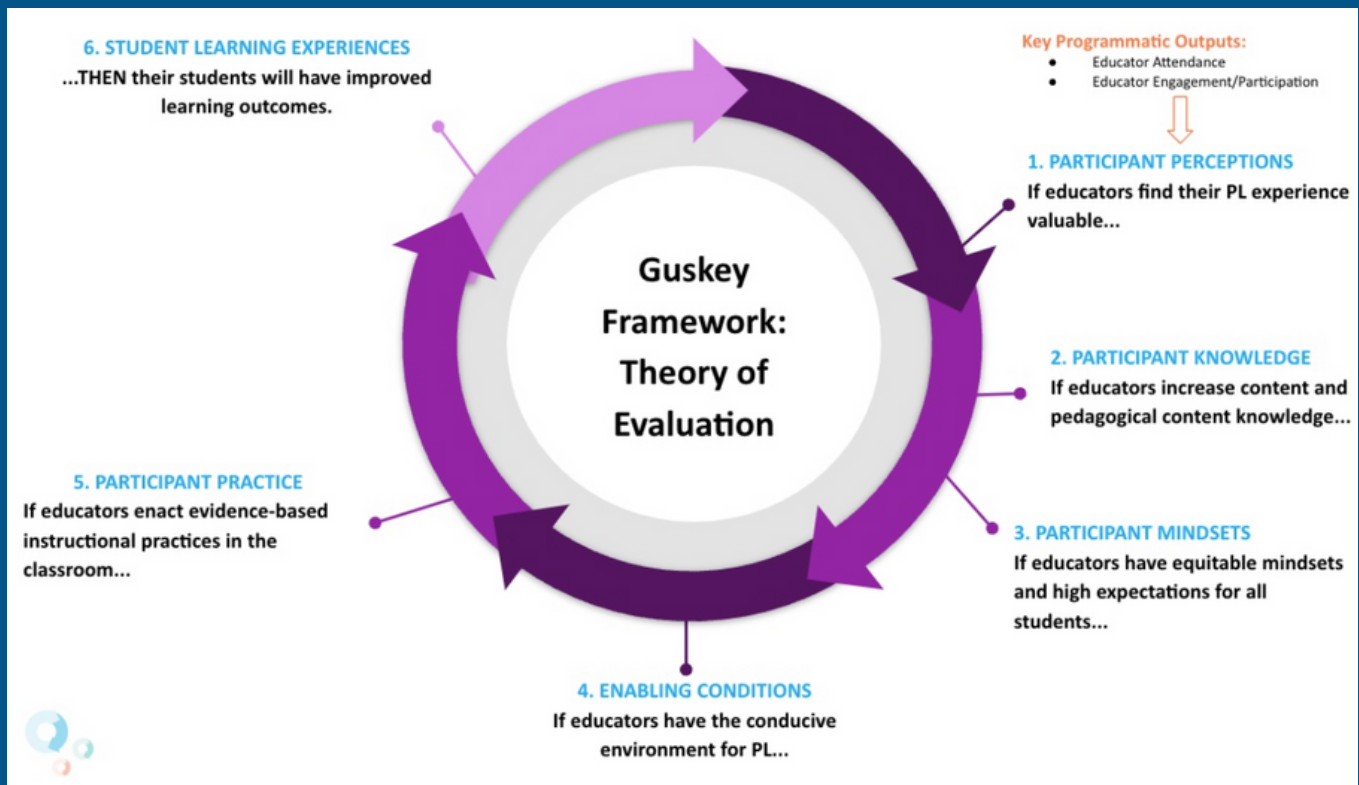
Measurement cycles at PL organizations vary depending on the scope of work, specifically the types of services and the timeline of partnerships.

Many PL providers, such as Teaching Lab (TL), collect data on two broad schedules.

- Targeted, short-term outcomes are gauged through ongoing data collection aligned to the implementation of specific PL services; this typically includes participant feedback surveys and course assessments that determine any pedagogical content knowledge gain and/or changes in self-efficacy related to specific PL course content.

- Medium- to long-term changes are tracked by collecting data at the beginning and end of the school year and/or partnership through instruments such as teacher surveys that determine any changes in mindsets and enabling conditions; classroom observations that assess teacher practices; and student surveys, student work samples, and analysis of district data that measure student outcomes.

With this in mind, TL's theory of evaluation adapted Guskey and the cadence of measurement (seen on the next page).

---

[1] Guskey, T. R. (2016). Gauge impact with 5 levels of data. Journal of Staff Development, 37(1), 32-37.

Guskey Framework: Theory of Evaluation

6. STUDENT LEARNING EXPERIENCES
...THEN their students will have improved learning outcomes.

Key Programmatic Outputs:
- Educator Attendance
- Educator Engagement/Participation

1. PARTICIPANT PERCEPTIONS
If educators find their PL experience valuable...

2. PARTICIPANT KNOWLEDGE
If educators increase content and pedagogical content knowledge...

5. PARTICIPANT PRACTICE
If educators enact evidence-based instructional practices in the classroom...

3. PARTICIPANT MINDSETS
If educators have equitable mindsets and high expectations for all students...

4. ENABLING CONDITIONS
If educators have the conducive environment for PL...

	Participant Perceptions	Participant Knowledge	Participant Mindsets	Enabling Conditions	Participant Practice	Student Learning Experiences
What we're measuring	Feedback on PL facilitation Perception of PL quality, acceptability, relevance, and applicability Net Promoter Score	Content, pedagogical content knowledge Curricular knowledge	Equitable mindsets and high expectations and beliefs	Peer trust, connection, and social capital Teacher ownership of PL Lab Leaders lead learning	Alignment of teachers' instructional practices to shifts and standards Self-reported and and student reported CRSE practices	Perceptions of learning environment Performance on tasks and formative and summative assessments
How we're measuring	End-of-Session /End of Course participant survey Ongoing Coaching Survey / End of Coaching Survey	Pre/post course knowledge assessment	Twice-a-year educator surveys	Twice-a-year educator surveys	Classroom observations Teacher evaluation results Twice-a-year student surveys	Student work samples Formative and summative assessments Twice-a-year student surveys

Level 1: Participant perceptions

TL administers ongoing feedback surveys on the facilitation and the overall quality and content of PL sessions to make adjustments from session to session. They also collect data on whether participants believe the PL is high quality, engaging, relevant, applicable, and feasible; whether they would recommend it to others (Net Promoter Score); and targeted feedback on the PL design and delivery to make any necessary changes to courses or coaching.

Level 2: Participant knowledge and mindsets

During the first session of a PL course, participants complete a short knowledge assessment consisting of multiple-choice quiz questions about the specific content and pedagogical content knowledge addressed in the sequence of learning for a content area. The results are used to adjust facilitation as needed. During the last session, participants complete the same set of multiple-choice quiz questions. This information is used to track how participant knowledge changes. For PL courses that are less focused on knowledge-building and more focused on skill-building or shifting practices, TL administers self-reported assessments related to teacher practice and/or self-efficacy.

In order to track how participant mindsets change throughout their partnerships, TL also collects data on key teacher mindsets that are predictive of student learning outcomes, such as the recognition of race and culture, high expectations that all students can and will learn, and growth mindsets in the Diagnostic Educator Survey and Follow-up Educator Survey.

Level 3: Enabling Conditions – Supportive structures and environment

TL's grassroots model of professional learning leans on the core belief that teachers deserve to feel motivated and supported by their peers to learn and grow. In teacher-led communities, educators are more likely to buy into their own development and work collaboratively with their colleagues to improve instruction. In the Diagnostic Educator Survey and Follow-up Educator Survey, TL asks teachers about their trust of other teachers, the level of collaboration they have with others in their schools, and support they receive from school leaders, which are parts of teacher social capital.

Level 4: Participant practices

TL conducts multiple classroom observations of teachers receiving coaching support to determine standards- and shifts-aligned instruction, using the core actions from the Instructional Practice Guides (IPGs) that are most emphasized in PL learning sequences (e.g., citing relevant evidence, productive struggle, student talk) throughout the partnership.

Level 5: Student learning experiences

The ultimate goal of TL's PL is to increase student learning and achieve educational equity. The twice-a-year student survey administered to students who have received TL's PL measures key dimensions of student learning experiences that are within teachers' control and highly predictive of student learning outcomes, including student-teacher relationships, self-efficacy, growth mindset, happiness and sense of belonging, being challenged, and culturally responsive teaching practices. TL also collects and analyzes student work samples at the beginning and end of some partnerships to track the quality of student tasks (i.e., rigorous, grade-level appropriate) and student performance on them. Finally, when available, TL analyzes formative and summative assessment data, comparing outcomes of classrooms and/or schools supported by TL to classrooms and/or schools not supported by TL.

———

Guskey is rooted in the assumption that success at one level leads to success at subsequent levels; yet, there is much we still have to learn about how different levels interact with one another. For example, must teachers highly rate their PL experience for it to positively impact their learning? Is organizational support and change a prerequisite for changes in teacher practice? Do mindsets change before practices change, or do practices change before mindsets change? Furthermore, while the Guskey Framework provides broad guidance on what and how to measure pre/post changes for PL participants and their students, rigorous evaluation requires a research design that includes meaningful comparison groups, randomized or otherwise, to truly determine the impact of any given PL program.

UnboundEd

Establishing the Reliability and Validity of Self–Created Tools

By Aliza Husain

In January 2022, Northwest Evaluation Association (NWEA) partnered with CORE Learning, a subsidiary of UnboundEd, to offer collaborative professional learning to educators with one of NWEA's large school system clients in Alaska. The program was designed as a three-credit college course and included synchronous professional learning and asynchronous coaching from NWEA and CORE's Online Elementary Reading Academy (OERA) for approximately 100 teachers. To assess changes in teacher knowledge, NWEA and CORE collected responses from approximately 100 teachers before and after they completed the course. An end-of-program report measuring teacher knowledge using a modified version of the Teacher Knowledge of Early Literacy Skills (TKELS) showed improvements in teacher knowledge. TKELS was selected as an external instrument because it was developed as a third-party measure of the impact of an early literacy initiative and it measured a wide selection of early literacy knowledge in which NWEA's client was interested.

Before implementing the modified version of TKELS, NWEA and CORE considered how well the modified assessment aligned with the collaborative professional learning program they were delivering. Since it was used solely for internal purposes, the assessment was considered a close enough match to the goals of the program. CORE continued to utilize and iterate on the assessment after the partnership and considered the following concerns about the modified assessment's reliability and validity as it pertains specifically to the OERA:

- Is the assessment *actually* measuring the constructs the OERA course aims to teach?

- Is the assessment *consistently* measuring the constructs the OERA course aims to teach?

To address these concerns, CORE's research and evaluation team traced the assessment's origins. They found that originally, the Regional Educational Laboratory (REL) Southeast developed TKELS. In contrast to measures that focus on subsets of early literacy skills, REL Southeast designed the TKELS as a broad instrument to assess various aspects of early literacy teaching, knowledge, and application.[2] This was why it served a high utility in NWEA's partnership with CORE. However, as the CORE program team continued to apply the assessment specifically with CORE's clients, they removed some questions not covered by OERA and added several questions relevant to OERA's content.

Because these changes were made to the validated original assessment, the modified instrument required re-validation. Given the limited modifications to the original measure, UnboundEd/CORE's research and evaluation team did not see the need to re-establish face validity or criterion-related validity, considering the program's constraints and available data. However, they believed it was crucial to re-establish construct validity to ensure the modified pre- and post-assessments accurately measured their intended content.

The research and evaluation team (R&E team) assembled a panel of experts familiar with OERA's content and the assessment's target constructs. They developed a protocol to establish construct validity, which involved answering questions about the assessment's intent, constructs, wording, and relevance. The panel's responses guided modifications to the assessment.

The R&E team then created a protocol for the panel of experts to establish construct validity, shown on the page that follows.

2 Folsom, J. S., Smith, K. G., Burk, K., & Oakley, N. (2017). Educator outcomes associated with implementation of Mississippi's K–3 early literacy professional development initiative (REL 2017–270). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. Retrieved from http://ies.ed.gov/ncee/edlabs.

After identifying a group of experts, ask each expert to first individually and independently answer the following questions. Then convene the group and collectively arrive at a consensus for all the following questions:

- What does the new test intend to measure?
- What is the overarching construct?
- What are the subconstructs?
- Why were certain items removed from the original test?
- Does the removal of these items allow the new test to better measure the thing it aims to measure?
- Why was a new item added to the original test?
- Does the addition of this item allow the new test to better measure the thing it aims to measure?
- Does each question measure what it intends to measure?
- Is the wording clear?
- Could the wording be misinterpreted/misconstrued in any way?
- Is the item connected to the overarching construct of the test?
- Does the test (pre- and post-independently) ask all the questions necessary to measure what it intends to measure?
- Are there any missing areas?
- Are certain areas of knowledge over-represented?
- Are any areas of knowledge under-represented?

Once the assessment was finalized based on the panel's feedback, CORE's research and evaluation team ensured accurate citation and sourcing of the assessment. They concluded that since the TKELS assessment was originally published through Institute of Education Sciences (IES) via REL, it could be used in the public domain.

The new assessment now aligns better with the OERA course and measures the course's intended constructs. UnboundEd plans to continue to pilot the new assessment in fall 2023 and intends to re-establish reliability during implementation.

Adapting an existing tool saved considerable time and effort compared to developing a new measurement instrument. It underscores the importance of finding, adapting, and using validated measures to assess program effectiveness, even though it comes with some challenges and demands on resources. This effort demonstrates NWEA and UnboundEd's commitment to using validated instruments to assess program impacts. For UnboundEd/CORE, this experience resulted in the development of a protocol for establishing construct validity that can be applied to other instruments used throughout the organization. They hope that the field continues to develop relevant measures, facilitating their adaptation and use by professional learning organizations as needed.

City Teaching Alliance* and American Institutes for Research

Embedding Transformative Social and Emotional Learning into Measurement

By Viticia Thames of City Teaching Alliance and
Lisa Merrill of American Institutes for Research

*formerly known as Urban Teachers

City Teaching Alliance and American Institutes for Research would like to acknowledge Roxanne White, Taylor Carter, Kevin Simpson, and Anthony Bowden for their support in this work.

There is evidence that teachers may operationalize guidance for developing students' social and emotional learning (SEL) skills through a deficit lens when students have disabilities or are students of color. For example, Kaler-Jones (2020) cited an instance when classrooms in a predominantly Black and Brown school had posters on the wall that defined social awareness as "keep your hands to yourself."[3] In response to similarly reported (mis)uses of SEL with students with disabilities and students of color, the Collaborative for Academic, Social, and Emotional Learning (CASEL) refined its definition of high-quality systemic SEL to include "a specific form of SEL implementation that concentrates SEL practice on transforming inequitable settings and systems, and promoting justice-oriented civic engagement—which City Teaching Alliance calls *transformative* SEL.[4] However, to date, there have yet to be research-based measures of transformative SEL (T-SEL) and its implementation in the literature.

> So while City Teaching Alliance had developed a T–SEL curriculum and T–SEL classroom observation rubrics, there were no ready–to–use, off–the–shelf teacher and student surveys aligned to teachers' and students' experience that City Teaching Alliance's external research partner could use to evaluate its effectiveness.

City Teaching Alliance outlined a measurable pathway analysis for why T-SEL PL would improve students' academic outcomes. In particular, their theory of action noted that if teachers used T-SEL practices with students, then students would experience greater psychological safety for taking the intellectual risks necessary for learning challenging content, which could then lead to improved learning outcomes. They then deduced that they would need to look to ongoing work in culturally responsive-sustaining education (CRSE) more broadly in order to identify student surveys that could measure something approximating students' psychological safety from stereotype threat as a learner of challenging content, and teacher surveys that approximated identity-affirming teaching practices.

After conducting a review of CRSE teacher and student surveys and reaching out to their professional networks, City Teaching Alliance learned about formative findings from Teaching Matters, a nonprofit based in New York and a fellow RPPL affiliate, that used teacher and student surveys aligned to New York State's CRSE framework.[5] As a member of the Bill and Melinda Gates Foundation's Networks for School Improvement (NSI), Teaching Matters' formative assessments identified Panorama student survey items aligned to the New York State Department of Education's CRSE framework that had statistically significant correlations to higher reading achievement.[6] As a result, City Teaching Alliance

---

[3] Kaler-Jones, C. (2020). When SEL is used as another form of policing. Medium. Retrieved September 22, 2023, from https://medium.com/@justschools/when-sel-is-used-as-another-form-of-policing-fa53cf85dce4
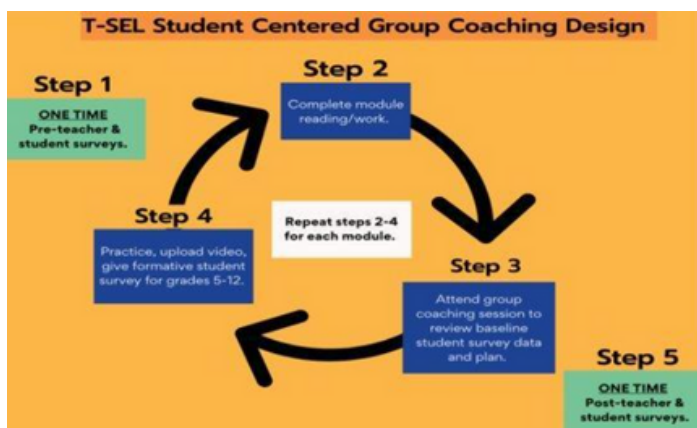
[4] Jagers, R. J., Skoog-Hoffman, A., Barthelus, B., & Schlund, J. (2021). Transformative social emotional learning: In pursuit of educational equity and excellence. American Educator, 45(2), 12-39.

[5] New York State Education Department. (2018). Culturally responsive-sustaining education framework. New York State Education Department. Retrieved September 22, 2023, from https://www.nysed.gov/sites/default/files/programs/crs/culturally-responsive-sustaining-education-framework.pdf

[6] Bell, J., & Istranyi, M. (n.d.). The impact of focusing on student identity. Teaching Matters. Retrieved September 22, 2023, from https://teachingmatters.org/insight/the-impact-of-focusing-on-student-identity/

decided to use the specific student Panorama survey items[7] that Teaching Matters had found were associated with higher reading outcomes as an objective summative measure. This also seemed like a promising pre- and post-summative measure, since the Panorama survey items were nationally normed. City Teaching Alliance could use the identified survey items across all four of their sites: DC, Dallas Fort Worth, Baltimore, and Philadelphia.

City Teaching Alliance also adopted Teaching Matters' approach of using student surveys for continuous improvement through using student surveys to inform iterative cycles of student voice-centered coaching. This was done by creating survey items aligned to what they believed students would experience if their teacher implemented practices described and demonstrated in each of the five asynchronous T-SEL Curriculum Modules that City Teaching Alliance Fellows took between group coaching sessions.

To understand how teachers experienced using these survey tools for their T-SEL practice with students, City Teaching Alliance partnered with AIR to interview teachers who participated in the continuous improvement PL about the program. Interviews explored their perspectives on administering the student survey, analyzing the results, and using the findings to improve their practice. Based on the interviews, AIR found that even though the T-SEL survey was short, concise, and aligned with the T-SEL PL, City Teaching Alliance struggled to garner buy-in from teachers to administer the survey. Some teachers thought the survey helped track their progress on T-SEL constructs like feelings of belonging, while others thought the surveys were too formal and preferred to talk to students about T-SEL.

A teacher who found the data useful to check in on how her classroom was forming said:

> "It helped me to see just based off of the certain questions the data was asking, I was able to be like, 'I expected that. I expected that to be high,' or, 'I expected this to be low. We were working on this more than that.' So it helped me gauge my classroom community as a whole and see different things grow as well or different things change based off of what we begin to talk about."

Other teachers did not buy in to collecting survey data and opted to have conversations with their students. Without the teacher's buy-in, the students also did not take the survey seriously.

> "I really did not use that survey data because, as I said, my students prefer to give me verbal responses and stuff. So we talked about it in class and I just let them know, 'Hey, I'm doing this course, this PD about T–SEL, this is what T–SEL is,' and I just asked them questions about what kind of things that they would be interested in learning about as far as this area. I gave them some examples and we talked about that and what skills they would like to develop. So, we had a discussion and I more so based my plans on the discussion. And also to be perfectly honest, some of them just straight out told me, 'I'm not reading these questions, Ms., I'm just doing it because you said you have to do it. So I'm just answering Strongly Agree to everything and not reading and submitting.'"

---

7 Mercer-Golden, Z. (n.d.). Rigorous expectations for student success. Panorama Education. Retrieved September 22, 2023, from https://www.panoramaed.com/blog/rigorous-expectations-student-success

Throughout the PL experience, some teachers believed the data was valuable to inform practice and the continuous improvement process. However, others were skeptical about the data collection—either that the results wouldn't be accurate or that they wouldn't tell them something they couldn't get from a conversation. Without teacher buy-in into the survey administration and data use process, gathering high-quality data over time is challenging. AIR offered substantial monetary incentives for conducting the survey, but many teachers still chose not to administer it to their students. If City Teaching Alliance continues this model of PL, they will need to find a way to create a community of teachers that understands the value of the survey and its results together with verbal or less formal student check-ins.

Teach For America
Addressing Internal and External Capacity Issues
By Katie Buckley

Teach For America (TFA), which recruits and develops a diverse corps of leaders who make an initial two-year commitment to teach in high-need schools, created a **measurement instrument that allows them to assess how teacher training on students' experiences of learning conditions is associated with improved educational outcomes.**

TFA began exploring Cultivate for Coaches (Cultivate), a survey and framework developed by Dr. Camille Farrington and the University of Chicago Consortium on School Research (Consortium)[8], because it aligns with the TFA's coaching model as it requires embedded coaching support for individual teachers (1:1 or in professional learning communities) using actionable tools and strategies. Cultivate is also aligned with a progress monitoring instrument, Elevate by PERTS, allowing teachers to monitor progress between survey administrations as they test new methods and techniques.

Following positive results from a small pilot, a team at TFA established a research-practice partnership with the Consortium to fully embed Cultivate into the work of TFA. This partnership was centered on the conviction that Cultivate's year-long system of support cannot simply be an add-on to TFA's programming but instead needs to be authentically embedded and aligned across TFA's system of supports, including current programming for corps members (CMs) delivered during pre-service (the summer before entering the school year). TFA worked with partners to co-develop training and support for CMs and coaches and align their support from pre-service to in-service.

TFA is now in the process of scaling Cultivate to their 38 regions in SY23-24. **A primary challenge to scaling Cultivate is capacity constraints for coaches and CMs in 5th-12th grade classrooms to embed it within their practice.**

TFA coaches
As the program's name, "Cultivate for Coaches," suggests, this program is intended to be driven by coaches. That is, TFA coaches must receive training on the survey as well as the research behind it, be equipped to support CMs with survey administration, ensure school leader buy-in as necessary, help their CMs unpack and understand their reports, set up individual and group sessions to drive changes in practice, and support their CMs in reflecting on the data in the spring. To alleviate some of the capacity constraints placed on their coaches, TFA sought to re-imagine their orientation and approach to classroom outcomes beyond academics by centering student perceptions of their classroom experiences. In doing so, TFA was able to weave Cultivate into core coaching moments, foundational coaching tools, and the CM Learning Plan, enabling CMs and coaches to have consistent opportunities to reflect on, identify, and support aligned changes in CM practice. Additionally, TFA developed a programmatic arc that includes Cultivate so that coaches could further understand what would happen with the survey and when, and plan for it, supporting them in fully integrating Cultivate into the overall CM experience.

TFA CMs
TFA asks CMs to administer Cultivate twice a year, along with Elevate multiple times a year, and use the data to drive changes in practice. This requires: (a) participation in survey instrument training to understand administration and reporting; (b) learning about the importance of classroom conditions for social, emotional, and academic development; (c) meeting with their coaches to prioritize next steps; (d)
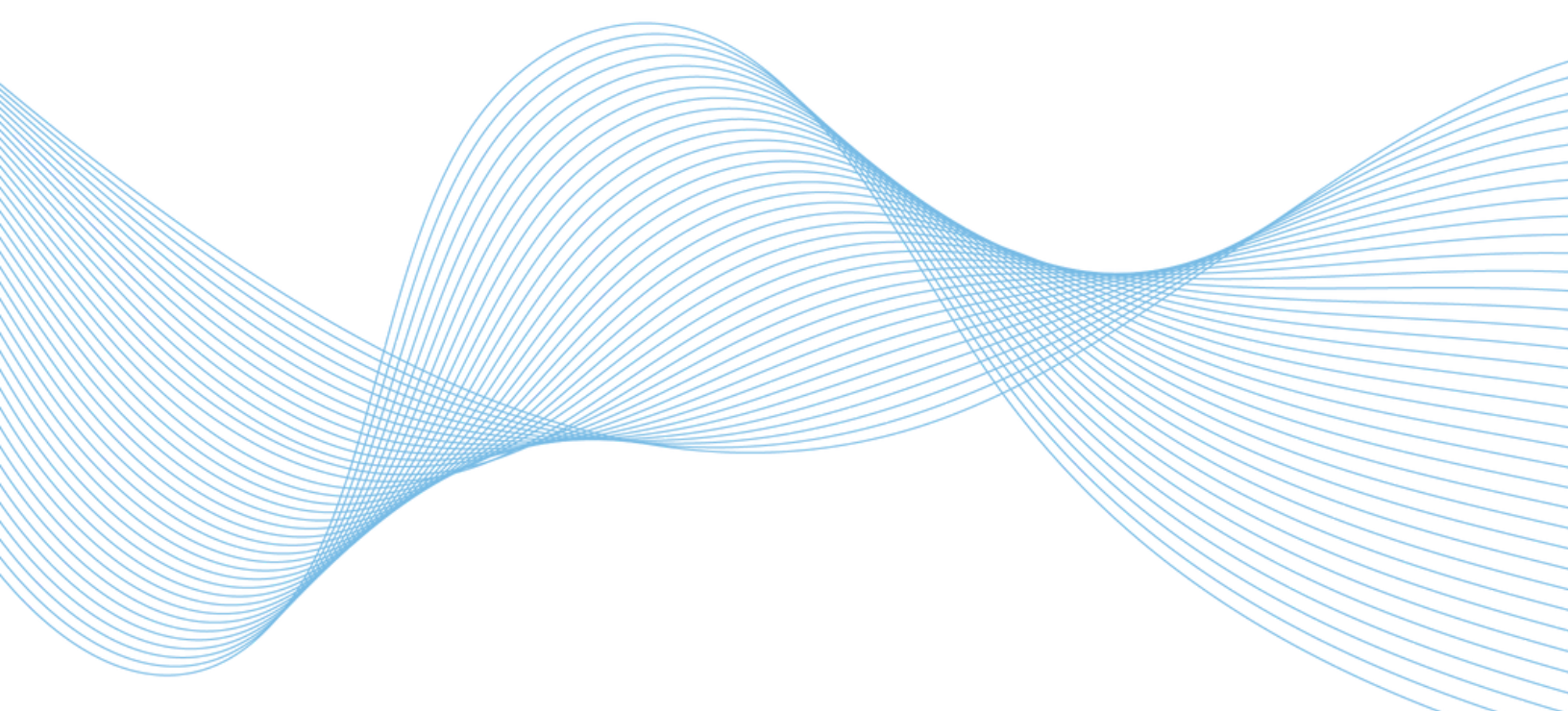
8 Farrington, C. A., Porter, S., & Klugman, J. (2019). Do classroom environments matter for noncognitive aspects of student performance and students' course grades? UChicago Consortium of School Research Working Paper.

internalizing Cultivate tools and resources to try out new strategies in the classroom; (e) using a progress monitoring tool (Elevate) to inform continuous improvement; and (f) reflecting on changes in their practice over the school year. This is on top of being a first- or second-year teacher and placed in under-resourced and high-need schools. To alleviate some of the capacity constraints on CMs, TFA has fully embedded Cultivate throughout the CM experience. Rather than having Cultivate as an "add-on" to their program, which is often the case with new initiatives, CMs' development is grounded in eliciting students' voices to improve classroom conditions for social, emotional, and academic development. This development begins at pre-service, and throughout in-service teaching, strategic opportunities and support are in place via coaching, learning, and professional learning communities.

To further alleviate capacity constraints among CMs and coaches, TFA sought to embed Cultivate across staff roles and responsibilities, training their staff to implement Cultivate, similar to how a district would work to embed a new professional development training across schools. To do this, TFA developed onboarding and training materials to ensure all CM-facing staff understand Cultivate and their role in supporting it, from Data Sharing Agreement (DSA) execution to survey administration operations to survey response tracking to coach and CM training. By building Cultivate into systems, processes, and infrastructure, automating functions where possible, and having national staff take over the administrative and operational aspects, TFA hopes to reduce the capacity constraints of CMs and coaches while increasing their engagement and ownership with the Cultivate instruments and resources.

While TFA still has far to go, they are pushing towards full scale implementation of Cultivate. For TFA, this doesn't just mean increasing the number of CMs participating across regions; it means organization-wide engagement, buy-in, and accountability that is indicative of authentic implementation and commitment to equitable classroom conditions. As TFA scales Cultivate's implementation, they are committed to sharing what they learn with the field, including implementation challenges and successes, student experiences in CMs' classrooms, and progress toward ensuring every CM is trained in creating equitable classroom environments for each student in their classroom.

Leading Educators
Supporting Systemic Coherence and Impact
By Laura Meili & Tim Tasker

Leading Educators

Leading Educators (LE) exists to ensure excellent and equitable teaching for all students. Successful implementation of **PL systems focused on equity and excellence requires considerable alignment and coherence within and across schools,**[9] which Leading Educators defines as **a set of enabling conditions: a clear, widely-held vision for high-quality teaching and learning, curriculum and assessment materials aligned to that vision, skilled instructional leaders to guide the learning, resources (e.g., time) to support it, and the data to understand its impact and to make adjustments.**

Few districts have yet had the support necessary to achieve consistent levels of coherence by putting these conditions in place,[10] which threatens effective implementation of collaborative professional learning. Additionally, focusing on these conditions at the school and district levels is a way of ensuring alignment and equitable resourcing across schools while also creating the opportunity for teachers and leaders to learn from each other through a collectivist (systems) approach rather than individualistic (personal) approaches to improvement.

Through the evolution of its own program, LE realized the need to codify and describe a specific set of enabling conditions. Originally, the organization launched as a fellowship model, where they trained teacher leaders outside their schools and hoped they would be able to transfer their learning back into their classrooms and schools. However, the barriers and challenges these teacher leaders encountered while attempting to do so illuminated the lack of coherence in their systems and pointed out the specific enabling conditions that were either underdeveloped or missing. Through these insights, LE shifted to a partnership model that directs coordinated supports and services across multiple levels of a school system, all focused on building coherence and developing the conditions for implementing high-quality PL for instructional improvement.

With considerable input from external experts and members of Leading Educators' teams, they created the school conditions and system conditions frameworks.
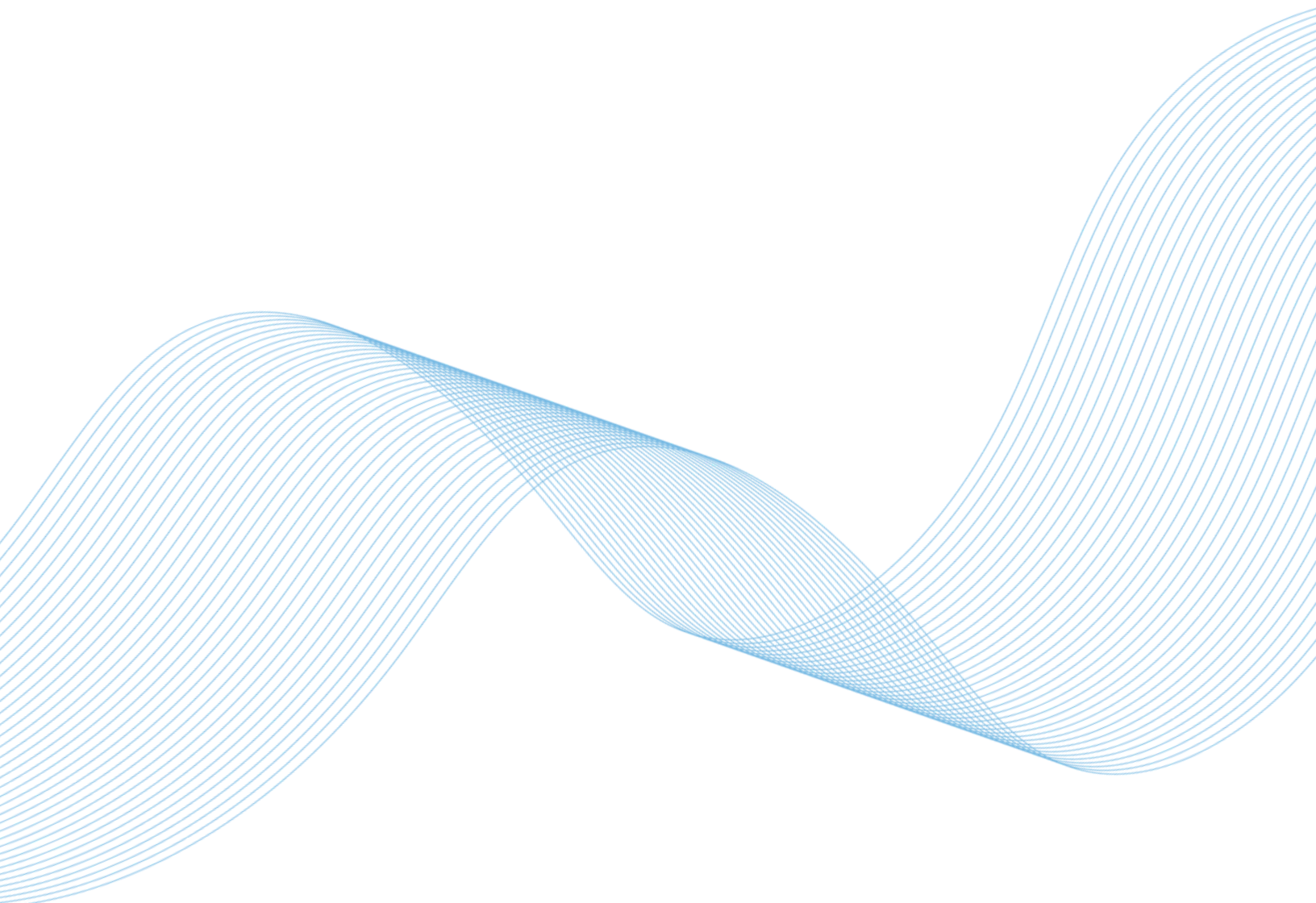
LE now administers the school and/or system conditions assessments twice a year in each of their partnerships, depending on each project's scope of work and the specific levels at which their work is focused. Stakeholders in multiple roles within a school or system complete the conditions assessment independently. LE aggregates their responses and shares a report back with partners to show how stakeholders perceive the development of each of six conditions. LE also triangulates across their staff's perspective and a set of artifacts in order to build consensus about the status of each condition and to plan actions that will further develop them.

The school and system conditions frameworks have given LE's partners a common language for understanding the key structures and systems that support effective PL for instructional improvement, clarifying their role as leaders and helping to define what success looks like in their work. Gathering data on the conditions has also helped partners see areas of progress to celebrate, places where conditions require further development to learn from, and priority areas for adjustments, where necessary.

9 Togneri, W., & Anderson, S. E. (2003). How high poverty districts improve. Leadership, 33(1), 12-16.
10 Fullan, M., & Quinn, J. (2016). Coherence making. School Administrator, 73(6), 30-34.

Nearly all of LE's partners significantly improve their conditions over the course of their partnerships, resulting in more consistent, effective, and sustainable PL. Many partners have embraced the conditions framework. In fact, several of the districts LE partners with now write goals aligned to their conditions. Moreover, the school and system conditions frameworks have also created internal coherence at LE, helping team members working across diverse projects to see connections and commonalities in the work they do and ensuring all staff understand and can describe how Leading Educators' theory of change plays out in practice to improve outcomes for leaders, teachers, and students.

Student Achievement Partners

Prioritizing Equitable and Culturally Responsive Approaches to Measurement

By Diana Cordova–Cobo

STUDENT ACHIEVEMENT PARTNERS

Note to our readers: SAP is not a "PL provider" like our other working group organizations; they are a systems provider. SAP works at the systems level with districts, agencies, other nonprofits, etc., which impacts their approach to evaluation. RPPL invited SAP to join our working group to highlight their work incorporating equitable and culturally responsive evaluation methods into how they measure the impact of their work and how they operationalize the shifts in their organization's grounding vision/mission related to equity when defining and measuring impact.

Historically, Student Achievement Partners (SAP) has been deeply committed to ensuring that all students, no matter who they are or where they live, are supported to access and successfully engage with grade-level literacy and mathematics content in the classroom. This commitment has resulted in the creation of resources like the Instructional Practice Guide (IPG) and Instructional Materials Evaluation Tool (IMET), which have been used nationwide by educators, systems leaders, and other nonprofits for over a decade.

SAP has now built on this foundation to help educators design instruction that leverages the assets and honors the brilliance of students historically underserved and marginalized by our education system. This next phase of work is being defined by their e² Instructional Practice Framework™,[11] which redefines high-quality instruction as being on grade level, joyful, culturally responsive-sustaining, and linguistically sustaining. The e² Framework is as much an internal roadmap for their work as it is an external resource for educators and was the culmination of years of their team's internal reflection and learning. During that period, it also became evident to SAP that refocusing their work's content alone was insufficient.

They needed to redesign how they engaged in project work with partners and how they measured the impact of their work in the education field.

For the latter, SAP turned to methodological and measurement approaches that aligned with their mission and ensured they were engaging educators and communities in the same ways they envisioned educators engaging with students—leveraging their assets and honoring their communities. **Following their pre-established approach to project work, SAP interviewed students, caregivers, and educators to understand how they defined success broadly and specifically in the context of SAP's work.** SAP also interviewed their staff to understand the impact they wanted to have on their project work, how that reflected the organizational statement of impact and progress measures, and what challenges and opportunities existed at the time for measuring the impact of their work. SAP then sought to learn from approaches to evaluation that aligned with their general equity-focused orientation to project work—namely, the Culturally Responsive Evaluation[12] and the Equitable Evaluation Framework[13]—and spent time orienting all their staff to this way of designing evaluation plans, regardless of their role.

---

[11] Arcos, S., Beltramini, J., Cordova-Cobo, D., & Swanson, C. (2023, July 7). The e2 Instructional Practice Framework: Toward a vision for high-quality instruction. Achieve the Core. Retrieved September 21, 2023, from https://achievethecore.org/peersandpedagogy/the-e2-instructional-practice-framework-toward-a-vision-for-high-quality-instruction/

[12] Hood, S., Hopson, R. K., & Kirkhart, K. E. (2015). Culturally responsive evaluation. In Newcomer, K. E., Hatry, H. P. & J. S. Wholey (Eds.), Handbook of practical program evaluation (pp. 281-317). Jossey-Bass.

[13] Dean-Coffey, J., Casey, J., & Caldwell, L. D. (2014). Raising the bar–integrating cultural competence and equity: Equitable evaluation. The Foundation Review, 6(2), 81-94.

Using all of these inputs, SAP established that its measurement approach:

- Is designed in service of educational equity and their organizational learning to fulfill their mission;
- Accounts for and seeks to reduce how power and inequality show up in education research, practices, and policies;
- Centers individuals and communities most proximate to the issue when considering evaluation design and implementation;
- Encourages authentic collaborations with stakeholder communities, where stakeholders and SAP staff work in partnership;
- Focuses on the evaluation process, as that is as important as the findings from the evaluation; and
- Assures evaluation findings are shared in ways that are meaningful to stakeholders and legitimize community knowledge and ways of knowing.

Once this foundation was established, SAP started working to create and curate the tools they would use to measure the impact of their work and, importantly, the process by which those tools would be employed throughout a project. Tangibly, this means having banks of interviews, focus groups, and survey questions instead of one static instrument, as well as having flexible protocols for artifact analyses or member checks and observation tools that can be adapted for various contexts and project foci.

Guided by their measurement approach principles, SAP also created a suite of tools for external use that align with the e² Instructional Practice Framework™ and emphasize process as much as outcomes. The e² Instructional Practice Suite,™ for instance, intentionally attends to gathering input from teachers and students, among other stakeholders, in a way that leverages their assets and honors their brilliance, through e²-aligned surveys and interview question guides. These qualitative inputs are vital to ensuring data reporting and decision-making processes include voices that may have been missing in the data collection process.

Although such an approach to measurement can require more time and capacity to enact, SAP has received affirming feedback from their internal staff and external partners that this approach yields meaningful data for their own learning and collective, coherent decision-making more broadly.

New Teacher Center
The Learning Environment: Measuring Beyond Teachers' Instructional Practice
By Lisa Schmitt

NEW TEACHER CENTER

New Teacher Center (NTC) recognized a need to supplement existing tools, like the Instructional Practice Guide (IPG) or the NTC-enhanced version created in partnership with Student Achievement Partners (SAP), with **observation indicators that would address classroom conditions to support social and emotional learning (SEL) and learning differences**. In 2017, NTC convened a panel of leading researchers/experts in SEL to participate in a series of conversations over several months to articulate the conditions that constitute an Optimal Learning Environment and observable indicators of those conditions. Together, NTC crafted a classroom observation rubric piloted in several U.S. sites. NTC then sought input from additional expert reviewers with specialties in SEL, learning differences, and Mind Brain Education (MBE) research.

The rubric was further refined and launched nationally in 2018 with partners who use NTC tools. The rubric includes indicators of teaching practices, student actions, and classroom interactions, and NTC crosswalked it with most major observation tools (e.g., Danielson Framework for Teaching) to demonstrate alignments that exist. Due to project delays, the rubric's psychometric properties have not been established, but NTC welcomes the opportunity to examine the reliability and validity of the instrument.

The framework helped convey NTC's belief that high-quality instruction is more than effectively implementing a rigorous curriculum and standards. It acknowledged the importance of establishing conditions that support students accessing and engaging in learning. NTC has heard from practitioners (instructional coaches, mentors, and teachers) who appreciate having the rubric descriptors that help them better understand what exemplary practice should look like, particularly in areas that may be less familiar to them. It helps to cultivate common understanding and reflects the best thinking of educational researchers about what high-quality classroom environments should include.

Practitioners also appreciate having indicators related to aspects of the instructional environment that were not included in the NTC IPG-based tool. Today, even when SAP's IPG tool is used, observers sometimes supplement with additional indicators from NTC's rubric to obtain a broader understanding of the classroom environment. The challenges associated with the expanded tool are due to its length. To ensure feasibility, the instrument is not intended to be used in its entirety; rather, mentors and coaches must be intentional when selecting a relevant subset of indicators to observe. This may result in limited opportunities to monitor improvements over time, depending on which indicators are observed in subsequent coaching cycles.

The resulting tool and framework demonstrated NTC's commitment to recognizing and supporting three interconnected domains of an Optimal Learning Environment:

- Create emotionally, intellectually, and physically safe environments.

- Implement equitable, culturally responsive, and standards-aligned curriculum and instruction.

- Meet the diverse needs of every learner.

NTC's forthcoming whitepaper will explore leveraging a decade of research to design instructional coaching for optimal learning–coming soon!