

SRI International

WestEd 

Implementing Classroom Observation Rubrics

**How are NGEI sites identifying and
using classroom observation
rubrics to prepare effective
teachers?**

Alix Gallagher
Melissa White
Naa Ammah-Tagoe
Ashley Boal
Ashley Campbell
Jaclyn Tejwani
Naomi Tyler

2017

© 2017 WestEd. All rights reserved. Permission to reproduce or adapt for non-commercial use, with attribution to WestEd and SRI International, is hereby granted.

WestEd is a research, development, and service agency whose mission is to promote excellence, achieve equity, and improve learning for children, youth, and adults. For more information about WestEd, visit <http://www.wested.org/>; call 415.565.3000 or, toll-free, (877)4-WestEd; or write: WestEd / 730 Harrison Street / San Francisco, CA 94107-1242.

SRI Education, a division of SRI International, is tackling the most complex issues in education to identify trends, understand outcomes, and guide policy and practice.

SRI International is a nonprofit research institute whose innovations have created new industries, extraordinary marketplace value, and lasting benefits to society. SRI International is a registered trademark and SRI Education is a trademark of SRI International. All other trademarks are the property of their respective owners.

This publication was made possible by a grant from the S. D. Bechtel, Jr. Foundation via its “Preparing a New Generation of Educators for California” initiative. Any opinions, findings, and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of the Foundation.

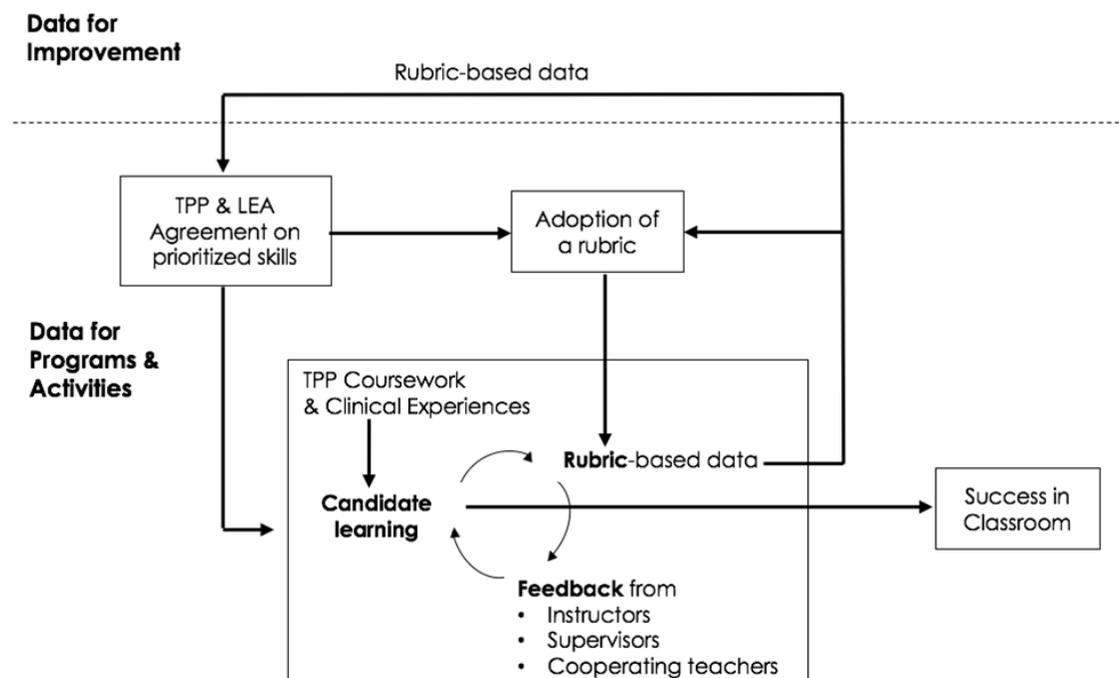
Suggested citation: Gallagher, A., White, M., Ammah-Tagoe, N., Boal, A., Campbell, A., Tejwani, J., & Tyler, N. (2017). *Implementing classroom observation rubrics: How are NGEI sites identifying and using classroom observation rubrics to prepare effective teachers?* San Francisco, CA: SRI International and WestEd.

Introduction

The New Generation of Educators Initiative (NGEI), funded by the S.D. Bechtel, Jr. Foundation, seeks to strengthen the current teacher preparation system in California so that new teachers enter the workforce prepared to implement Common Core State Standards (CCSS) and the Next Generation Science Standards (NGSS). The Foundation has developed a theory of action to guide reform that focuses on five Key Transformation Elements: partnership, prioritized skills, practice-based clinical preparation, formative feedback on prioritized skills, and data-driven continuous improvement.

WestEd and SRI International are conducting a formative evaluation to track NGEI implementation and outcomes at the 11 NGEI grantees (i.e., TPPs and their district partners) that received comprehensive grants in Phase 2. One of the core NGEI requirements is that each partnership (campus and district) identify prioritized skills and a classroom observation rubric to measure candidate progress towards those skills. This is because high-quality rubrics can play a central role in preparing effective teachers and supporting ongoing improvement of preparation programs (Exhibit 1).

Exhibit 1. Classroom Observation Rubrics at the Center of Candidate and Organizational Learning in NGEI



As the theory of action shows, classroom observation rubrics (hereafter “rubrics”) can operate at the center of individual and organizational learning. Rubric data from assessments of candidate practice can:

- Build consistency in formative assessment and feedback candidates receive about their strengths and areas to improve, thus *facilitating candidate learning*
- Be used to evaluate whether candidates have the skills necessary to be *recommended for a credential*
- Support *continuous programmatic learning* about trends in candidate performance and, therefore, aspects of the coursework and clinical experience that should be revised

In addition to generating valuable data, rubrics can play a more foundational role in NGEI partnerships. When campus-district partners collaboratively select or develop rubrics, the *rubrics articulate a consensus view of effective teaching*. Rubrics are then a powerful tool for communicating that vision of effective teaching to a range of stakeholders—professors, district administrators, university supervisors, cooperating teachers, and candidates. When used consistently they can break down gaps candidates might otherwise perceive between the theory taught in courses and the practice learned in clinical settings; they can also smooth the transition from preservice preparation into induction. The range of powerful uses for rubrics, however, adds to the complexity of selecting and using them. This Innovation Highlight is devoted to surfacing some of those complexities and then sharing some of the ways NGEI partnerships started working with rubrics during the 2016–17 school year. (Appendix A provides an overview of the rubrics selected or developed by NGEI partnerships.)

Central Issues in Using Rubrics as Tools for High-Quality Teacher Preparation

From both existing research and our site visits to all 11 grantees during spring 2017, we identified five issues that grantees might need to grapple with in order to use rubrics as a lever for successful teacher preparation. The discussion is not to advocate for any particular response to these issues, but rather to highlight them, and then describe particular approaches NGEI partnerships are taking in response to them.

1. Selecting the Focus

One of the key issues in using rubrics to measure instructional practice in a teacher preparation program is *deciding what dimensions of teaching are most important to measure for various purposes*.

I would say that [identifying hallmarks of effective teaching] is something that I can do now, because we've started to have these conversations, but if you would have asked me even a year ago I could have only been able to speak from [my personal] lens. Now I think we're all pretty bought in to the prioritized skills that we have identified on our observation tool in terms of looking for really strong instructional design from our candidates, looking for evidence that it's based on knowledge of students, ... knowledge of assessments, ... alignment to standards, and use of best practices in the field.

Clearly, as part of credentialing, programs need to comply with the California Commission on Teacher Credentialing's (CTC) requirement to ensure that all candidates are adequately prepared to meet the Teaching Performance Expectations (TPE). There are six TPEs which include a total of 45 nested elements. All TPEs and nested elements are specified as having comparable importance and value. Therefore, while there may be a tradeoff regarding breadth versus depth, there is a clear rationale for aligning ongoing formative assessment and feedback with the credential requirements.

Another approach would be to focus on a relatively small set of what Ball and Forzani (2011) call "high-leverage practices," defined as those teaching activities that are essential; "if [teachers] cannot discharge them competently, [they] are likely to face significant problems. Competent engagement in them would mean that teachers are well-equipped to develop other parts of their practice and become highly effective professionals." The theory is that by identifying a relatively narrow subset of high-leverage practices, programs could better ensure a focused

and coherent approach to how candidates are introduced to practice, and receive feedback on those prioritized skills.

2. Validity as Relative

Rubrics and the associated processes around their use can make them differentially “valid” for different purposes. Validity is about more than just accuracy; it is about how accurately the tool measures the construct it is designed to measure, *relative to the consequences* associated with the measurement. *To make valid, high-stakes decisions about candidates might require a more comprehensive rubric and more precise calibration process* than required to provide candidates formative feedback about specific skills or ascertain more broadly if candidates are typically developing high-priority skills as they progress through a program. This suggests the importance of grantees thinking about various ways to understand candidate skills and progress for various purposes during their teacher preparation program. As such, it is both complicated (there is no silver bullet rubric) and freeing (because there is no one right approach for all purposes).

3. Alignment between TPP and Partner Districts vs. Measuring Growth

Partnerships are best served by rubrics that can detect candidates’ progression as they develop skills through their program and also smooth the transition between preservice and induction. The ideal rubric would be fine-grained enough to capture the many steps of growth between beginning teacher performance during preservice clinical experiences and exceptional inservice teaching. However, such a rubric would require many score points, which would decrease the chances that it could be used reliably (because differences in levels would necessarily be small). On the other hand, a rubric with fewer score points (e.g., a 3-point rubric) is likely easier to calibrate on but it might be prone to other problems (e.g., the fact that score points are too far apart to capture the extent of growth candidates are likely to make during student teaching and the related pressures raters might feel to rate candidates a “1” in the fall and a “2” in the spring to represent candidate growth). This is a *measurement quandary, with all options requiring tradeoffs*.

4. Rolling out the Rubric to Various Stakeholders

For rubrics to function as teacher preparation improvement levers—communicating a vision of effective teaching and helping support and track individual and program-wide progress towards that vision—a wide range of stakeholders need to understand the rubric, and its purpose, at a high level. By “understand” we mean an awareness of the basic dimensions of the rubric and a general sense of how the rubric describes more and less effective instruction. Although district administrators would not use the rubric to rate candidates, their buy-in and understanding could support a culture where the rubric supports consistency in the messages candidates receive about the qualities of effective instruction. Figuring out how to stage an initial rollout and ensure sufficient ongoing training is an essential management task for NGEI leaders.

5. Getting Calibrated on Ratings and Feedback

Calibration is a level of knowledge about the rubric that goes beyond the understanding of the rubric that a broad range of stakeholders should have (as described above). Calibration involves training users to see instruction similarly when using the rubric’s lens such that they attend to the same aspects of instruction and arrive at the same conclusion about the level of proficiency evidenced by that sample of instruction. ***Rubrics can most effectively be used to support candidate learning if there is calibration on ratings as well as consensus around how feedback will connect to rubric dimensions.*** As an example, let’s imagine a rubric designed to measure candidate proficiency on TPE 1.5: “Promote students’ critical and creative thinking and analysis through activities that provide opportunities for inquiry, problem solving, responding to and framing meaningful questions, and reflection.” This TPE is clearly critical for effective teaching, but it is also very broad. Partnerships could try approaches to help raters calibrate on this dimension by creating “look fors” and undertaking a calibration process. Partnerships might also develop a framework around what it means to “frame meaningful questions” in various teaching and learning situations so that supervisors and cooperating teachers can clearly articulate a vision of effective questions and scaffold candidates’ skills development through their feedback. Further, engaging in calibration will allow for supervisors and cooperating teachers to provide consistent feedback to candidates, allowing candidates to improve their practice towards a clear, agreed-upon goal. Of course, building consensus and calibrating in this way takes time, and grantees need to

decide which stakeholder groups need to understand the rubric and which need to invest the additional time required to be calibrated.

Examples of NGEI Partners Work with Rubrics

Since the start of the second phase of NGEI grants, all 11 partnerships grappled in unique and important ways with rubrics as part of their reform work. Below we highlight aspects of seven approaches that we think would be instructive for the broader group. In doing so, we focus on one element of that partnership’s work. The descriptions do not line up in a one-to-one way with the issues listed above, although all overlap on one or more issues and/or exemplify some aspect of the theory of action. Additionally, our selection of these approaches does not imply that these are the “best” or “only” ways to make rubric work valuable. Rather, we hope that these descriptions stimulate thoughts within the NGEI Learning Community.

Box 1. Articulating a Shared Vision Among Key Stakeholders

Project leaders from CSU Sacramento (CSUS) and Sacramento City Unified School District (SCUSD) emphasized the importance of identifying shared priorities and selecting a rubric that reflects their deep partnership. Through small group work among the leadership, the team vetted and revised a list of skills that they value in teaching, from candidates to veterans. Having established a common language around excellent teaching, the team opted to use the district’s existing formative feedback tool as the rubric for this project. The rubric highlights a number of the prioritized skills and will provide the stakeholders—CSUS faculty and supervisors, and partner administrators and mentor teachers—with the tools to rework coursework and fieldwork, so as to provide candidates with a common discourse across the key elements of their teacher preparation program. Moreover, increased alignment between CSUS and SCUSD teaching expectations provides CSUS candidates seeking positions with SCUSD the possibility of consistent performance expectations through their induction years and into their teaching career in Sacramento. Through a focus on partnership development during this first year, CSUS and SCUSD defined the core values that will guide all of their work together.

Box 2. Customizing Existing Resources

The project leader from CSU Long Beach (CSULB) and various stakeholders from Long Beach Unified School District (LBUSD) collaborated to develop the rubric at the center of their NGEI grant work. As they worked together, the project leader from CSULB brought the perspective of the university and emphasized the importance of the California Teaching Performance Expectations (TPEs). The LBUSD stakeholders brought the perspective of the district and emphasized the importance of the California Standards for the Teaching Profession (CSTP) and LBUSD's induction rubric used to support beginning teachers' development. While the team considered expectations from both the TPEs and the CSTPs as a starting point, the team began the development process by reviewing the TPEs and determining which were most important to the priorities of the CSULB program. Through this process, the partners developed four TPE-based categories for inclusion in the rubric. After selecting the TPE-based areas of interest, the LBUSD stakeholders worked to map relevant CSTPs and induction rubric indicators onto the established categories and develop observable and measurable metrics for inclusion in the rubric. The end result is a rubric that builds from the TPEs, CSTPs, and district induction rubric but is customized to better align with the university and district priorities.

Box 3. Tracking Calibration

CSU Fullerton and their district partners identified and selected a math-specific rubric, the Mathematics Classroom Observation Protocol for Practices (MCOP2), to address the need for improved mathematics instruction. Thus far, university supervisors, clinical coaches, CSU Fullerton faculty, district staff (principals and curriculum specialists), mentor teachers, and math teachers have participated in training and calibration activities.

To ensure all stakeholders understood the context and relevance of the MCOP2 and to work toward calibration, participants attended an in-person training. During the training, CSU Fullerton highlighted new approaches to mathematics instruction (e.g., instruction focused on eliciting student thinking to drive learning and an emphasis on developing students' mathematical reasoning and sense making rather than teacher-delivered rote procedural knowledge), introduced the MCOP2, and gave participants opportunities to practice using the tool. After watching video examples, participants independently rated the video, discussed specific evidence for each rating, and adjusted their ratings when necessary to align with the common lens being established. After the training, faculty participants were required to complete a calibration homework assignment (optional for district participants). For the assignment, they viewed an additional video and submitted their scores to CSU Fullerton via a Google form. The associated Google sheet automatically tracks who has submitted scores for calibration and documents the ratings each person selected.

As multiple participants attend the training and complete the calibration homework assignment, this process will produce a running record of calibration data, allowing the project leadership to check consistency in ratings at any time and identify individuals who may need more support to achieve calibration. While this process is only in the beginning stages, the infrastructure established to collect calibration data will enable CSU Fullerton to carefully examine calibration over time. Using this process and infrastructure to conduct subsequent calibration checks may allow the project leadership to track the extent to which calibration shifts over time.

Box 4. Calibrating on a Shared Vision of Instruction

CSU Chico and their partners at Chico Unified School District (CUSD) collaboratively selected the TNTP Core rubric for a number of reasons. Chief among them was the alignment between the vision of effective teaching embedded in the rubric and CCSS/NGSS. Additionally, the TNTP Core has a manageable number of performance domains (four), including “academic ownership,” which focuses on the extent to which students are doing the thinking in the classroom. NGEI leaders see the rubric as a tool for articulating what this fundamental shift from traditional to CCSS-aligned instruction looks like in classrooms.

When they adopted TNTP, the partnership purchased a training and calibration program through which leaders from both the district and university got trained and calibrated. In addition to NGEI leaders from both partner organizations, a representative from the teachers' union as well as supervisors who work with candidates not participating in NGEI received training. This selection opens up possibilities for the rubric to spread into the district and across multiple credential programs during and after the grant. Finally, there is an option to tailor training and to explore using a train-the-trainer model in future years. Given the length and stringency of the training process, flexibility on how to conduct the training will be key in broadening the use of the rubric.

Box 5. Rubrics as Tools to Build Bridges: Connecting Coursework and Clinical Practice

CSU Monterey Bay (CSUMB) and Monterey Peninsula Unified School District (MPUSD) worked together to develop a rubric that describes effective STEM teaching. The rubric has roots in multiple sources, including the 5d+, existing CSUMB and MPUSD descriptions of effective teaching, and the TPEs. The rubric is designed to be developmental (meaning that it outlines a

progression from emerging to well-developed teaching skills) and to be "non-evaluative" (meaning that candidates' progress through the program is not determined by their ratings on this tool).

Candidates are introduced to the rubric in their science and/or math methods courses where, among other things, they use the rubric to formally review the STEM practices of teachers in the field. As a result of this exercise, feedback given to classmates on STEM lessons they plan and implement in class is also shaped by the tool. Candidates reported that referencing the rubric to give classmates feedback on their teaching and receiving feedback from classmates on their own rehearsal teaching helped them to internalize what effective teaching looked like and to see how to progressively become more effective. Because the rubric was developed jointly with the MPUSD and expresses a shared district-campus vision of effective teaching aligned with CCSS and NGSS, candidates reported that they found the expectations on the rubric matched the expectations their cooperating teachers and administrators have for their instruction.

Box 6. Rubrics as Tools to Build Bridges: Aligning Preservice and Inservice Feedback

Building on a strong Phase 1 partnership, CSU Stanislaus, Ceres Unified School District (CUSD), and Turlock Unified School District (TUSD) jointly selected the 5D+ Rubric for Instructional Growth and Teacher Evaluation (5D+ Rubric) as a tool that exemplifies the partnership's understanding of effective teaching for both preservice and inservice teachers. The 5D+ Rubric, designed by the University of Washington Center for Educational Leadership based on their 5 Dimensions of Teaching and Learning instructional framework, encompasses five instructional dimensions and a sixth dimension of professionalism. The developmental orientation of the rubric makes it appropriate for supporting teachers' growth while they are in the preparation program as well as after they are hired by the Ceres and Turlock districts.

The CSU Stanislaus/CUSD/TUSD partnership plans to use the 5D+ Rubric to serve as the foundational formative and summative feedback tool starting during candidates' student teaching and continuing through induction. In spring 2017, faculty, university supervisors, district administrators, and cooperating teachers received calibration training on the rubric and began using it as the basis for formative feedback in student teaching. With the goal of facilitating a smooth transition from preservice to inservice teaching, district induction leaders are developing plans to incorporate the 5D+ Rubric into induction programming beginning in fall 2017.

Because the partnership envisions using the 5D+ Rubric for formative feedback beyond preservice preparation, university and district stakeholders reached consensus that a reasonable rating on most rubric

indicators for a candidate during their student teaching would be “basic,” the second of four rating levels. Using the 5D+ Rubric in the induction program may provide a valuable opportunity for the partnership to home in on how teachers move from “basic” to “proficient” or even “distinguished” ratings early in their career.

Box 7. Scope and Sequence for Rubric Rollout

CSU Bakersfield (CSUB) and Bakersfield City School District (BCSD) agreed to assess resident progress in their program through the use of the Danielson Framework for Teaching, specifically focusing on the 10 observable components from Domain 2: Classroom Environment and Domain 3: Instruction of the Danielson. The BCSD Curriculum Coordinator first introduced mentors and residents to the Danielson Framework through an orientation to the rubric at the beginning of the 2016–17 school year. Since then, the partnership has hosted monthly meetings with mentors and residents to scope and sequence the rubric by diving deeply into specific components at each meeting through the use of Teachscape videos, evidence collection, and in-depth discussion. Early trainings focused on the components related to building rapport with students, creating a culture for learning, and classroom management. Later trainings discussed components related to student engagement, assessment, and questioning techniques. Residents are only rated by their mentors on the components they have received in-depth training on. Throughout the semester, as more Danielson components are added to their formal observations, residents develop a comprehensive understanding of the rubric, the expectations for teaching, and the accompanying feedback.

Discussion

We hope this Innovation Highlight underscored important issues in rubric use and illustrated a variety of approaches campuses and district partners are using as the NGEI rubrics work unfolds. Collectively, these approaches show that there is no single rubric that solves all problems; similarly, every rubric can be used in ways that benefit candidates and teacher preparation programs. However, ensuring those benefits will require substantial strategic decision-making and ongoing adjustments based on data. As the logic model shows, the success of rubrics should be evaluated relative to two main goals: (1) the extent to which they and the way they are used provide candidates feedback to support learning to teach; and (2) the extent to which they provide teacher preparation programs insights into how well

they are preparing candidates to demonstrate the high-priority skills they and their district partners believe are critical for success in the classroom.

Over the summer all grantees will face one more issue—ensuring that their rubrics are aligned with the new TPEs. Every rubric grantees selected can become part of a system that complies with the new CTC requirements. We encourage partnerships to discuss their approaches and learn from each other’s innovations as this and other issues arise in the future.

Reference

Ball, D. L., & Forzani, F. M. (2011). Building a Common Core for learning to teach: And connecting professional learning to practice. *American Educator*, 35(2), 17.

Appendix A. NGEI Classroom Observation Rubrics

Campus	Rubric name	Newly developed or preexisting rubric?
Bakersfield	Danielson's Framework for Teaching	<input type="checkbox"/> Developed ^a <input checked="" type="checkbox"/> Preexisting ^b
Channel Islands	California Standards for the Teaching Profession (CSTP)/San Francisco Unified School District (SFUSD) Observation Rubric	<input checked="" type="checkbox"/> Developed <input type="checkbox"/> Preexisting
Chico	TNTP Core Rubric	<input type="checkbox"/> Developed <input checked="" type="checkbox"/> Preexisting
Dominguez Hills	Assessment Summary of Teacher Performance	<input checked="" type="checkbox"/> Developed <input type="checkbox"/> Preexisting
Fresno	Fresno Region Common Rubric	<input checked="" type="checkbox"/> Developed <input type="checkbox"/> Preexisting
Fullerton	Mathematics Classroom Observation Protocol for Practices (MCOP2)	<input type="checkbox"/> Developed <input checked="" type="checkbox"/> Preexisting
Long Beach	California State University Long Beach Multiple Subjects Credential Program Observation Rubric	<input checked="" type="checkbox"/> Developed <input type="checkbox"/> Preexisting
Monterey Bay	STEM/CSTP Teaching Rubric	<input checked="" type="checkbox"/> Developed <input type="checkbox"/> Preexisting
Sacramento	Sacramento City Unified School District Observation and Coaching Tool (SCUSD OCT)	<input type="checkbox"/> Developed <input checked="" type="checkbox"/> Preexisting
San Luis Obispo	Teacher Candidate Observation Report and Feedback Form	<input checked="" type="checkbox"/> Developed <input type="checkbox"/> Preexisting
Stanislaus	5+ Dimensions of Teaching and Learning Rubric (http://info.k-12leadership.org/5d-teacher-evaluation-rubric)	<input type="checkbox"/> Developed <input checked="" type="checkbox"/> Preexisting

NOTE: We have included links to those observational rubrics that are publicly available or posted on the NGEI Learning Community site, <http://ngei.wikispaces.com/Grantee-developed+Resources>.

a. Developed through an NGEI partnership

b. Developed by outside entity.