



Kentucky Department of Education Professional Growth and Effectiveness System (PGES) Implementation Rubric





This rubric is being designed for district leaders to self-assess and understand how teacher and leader evaluation systems are being implemented across their district. It uses definitions and evidence of strong and weak implementation to help district leaders evaluate where roll-out has been strong to date and to help guide planning around next steps.

This toolkit builds on the Measures of Effective Teaching (MET) principles defined by the Bill and Melinda Gates Foundation as well as the work in Kentucky to create a robust system for Professional Growth and Effectiveness System (PGES). It is intended to be a resource for district leadership as they design and implement PGES locally, and will be used in collaboration with state-provided supports (e.g. Effectiveness Coaches and PGES Coaches, among other supports) to facilitate strong implementation, and, increase the likelihood of successfully increasing the number of highly effective educators across the state, and, ultimately, lead to improved student outcomes.




Questions, comments, and feedback about the rubric can be shared with your district and/or regional coach(es), Todd Baldwin, Executive Strategic Advisor, Office of Next Generation Learners, KDE (todd.baldwin@education.ky.gov), and/or Sara Kerr, Program Manager at EDI (skerr@deliveryinstitute.org).

Kentucky Department of Education Capacity Frameworkⁱ

KDE's Capacity Framework provides guidance around core behaviors and actions that, when applied with intentionality and fidelity, increase the likelihood of successful implementation. These capacities are embedded throughout the rubric, as they are important indicators of success and warrant self-reflection. They push us to think beyond the "what" to the "how" and "how well." Where applied, the reader will note one of four customized bullets next to a "Question to Consider" that aligns with the four core capacities outlined by the Kentucky Department of Education. Alignment is indicated, where applicable, by the presence of one of four symbols as denoted in the table below.

Core Capacity	Characteristics	Application
 Human	High-capacity individuals are strategically situated in the system, and there is evidence of sustainable human capacity growth.	Intellectual proficiency, will
 Organizational	Culture is shaped or evidenced by the degree to which interactions, collaborations, and communications are positive and change-oriented.	Interactions, collaborations, communications
 Material	Consider existing resources from within and outside system, as well as the allocation of those resources. Resources must be allocated according to strategic priorities.	Fiscal resources, material resources
 Structural	Problems arise when structures do not fit the organizational goals, or do not facilitate the change process. The quality of structures must also be considered.	Roles, procedures/routines, policies, hierarchies

Rubric

Category	Questions to Consider	Weak (1)	Strong (4)	Evidence to look for
Measure Effective Teaching				
1. Set expectations – Do stakeholders agree on a definition of effective teaching and what an effective evaluation system should look like?				
a. Clarity and Transparency	<ul style="list-style-type: none">  Is there a shared definition of effective teaching and what an effective teacher and leader evaluation system should look like?  Did your system seek system design input from key stakeholders at all levels? • Do key stakeholders broadly agree on the set of knowledge, skills, and behaviors that enable improved student learning? • Do stakeholders at all levels demonstrate a deep understanding of your evaluation system and its importance as a driver of improved teacher effectiveness and student learning? 	<ul style="list-style-type: none"> • No evidence that single definition of effective teaching exists. • System designed “in a vacuum” with little to no input from stakeholders; decisions made without consideration of stakeholder opinions. • Majority of stakeholders unable to articulate definition or describe major components of PGES. 	<ul style="list-style-type: none"> • A clear and shared definition of effective teaching has been established. • Input solicited from stakeholders at all levels and used authentically to inform PGES design and major decisions were made openly. • Stakeholders can articulate that definition and describe major components of the PGES. 	<ul style="list-style-type: none"> • Shared language around effective teaching exists at all levels, among district administration, school leaders, teachers, parents, etc. • When probed, stakeholders at all levels able to clearly articulate key components of PGES. • Feedback from the field suggests consensus around the characteristics of an evaluation system and their system’s design.
b. Engagement and Communication	<ul style="list-style-type: none">  Were a broad and inclusive set of 	<ul style="list-style-type: none"> • Evaluation system design and implementation was done with little or limited 	<ul style="list-style-type: none"> • Stakeholder opinions were sought throughout the entire 	<ul style="list-style-type: none"> • Attendance at stakeholder meetings represents a full and

Category	Questions to Consider	Weak (1)	Strong (4)	Evidence to look for
	<p>stakeholders involved in designing your evaluation system?</p> <p>💰 Have resources been allocated toward the development of a robust communications strategy?</p> <p>😊 Is there a single responsible owner for communicating about initiative and engaging stakeholders?</p> <p>📅 Are there routines in place to assess the district's engagement strategy and make mid-course corrections where necessary?</p> <p>👍 Are communications about PGES system design and implementation delivered in a clear and timely fashion?</p> <ul style="list-style-type: none"> Is there broad awareness of and buy-in into PGES among a diverse group of stakeholders at all levels? 	<p>involvement of key stakeholders.</p> <ul style="list-style-type: none"> Few or no resources have been allocated (from either new or repurposed funds) toward engagement/communications activities. There is a lack of single responsible person for communication/engagement work. Outreach is characterized by unclear or incomplete communications with very limited regard for audience. Awareness of and support for PGES is non-existent or limited; stakeholders are either uninformed, unsupportive, or both; resistance from the field is evident. 	<p>design and implementation process.</p> <ul style="list-style-type: none"> Communication and engagement efforts are adequately supported with funding and human capital support. District has assigned clear ownership for communication/engagement efforts. High-quality, tailored communications materials have been developed and widely distributed. Stakeholders demonstrate buy-in and actively push for strong implementation of PGES. 	<p>diverse set of players.</p> <ul style="list-style-type: none"> There is clear ownership for the development and implementation of a communications strategy. Stakeholders have comprehensive access to information about PGES via multiple modes (e.g. webinars, newsletters, guides, meetings, etc.) There is an existing structure and/or routine in place to ensure vision for effective teaching is diffused widely and frequently (e.g. as evidenced in CDIP)
<p>2. Use multiple measures – Does your district's Certified Evaluation Plan (CEP) provide clear guidance on the selection and integration of multiple sources of evidence to inform a teacher's summative performance rating?</p>				
	<p>Does your district's CEP set clear expectations for</p>	<ul style="list-style-type: none"> CEP does not clearly lay out expectations for the use of minimum 	<ul style="list-style-type: none"> CEP incorporates multiple sources of evidence, including, 	<ul style="list-style-type: none"> Sources of evidence are manageable from a school leader and

Category	Questions to Consider	Weak (1)	Strong (4)	Evidence to look for
	<p>evaluators' use of multiple sources of evidence teacher and leader performance, rather than relying on a single metric?</p> <ul style="list-style-type: none"> • Are the sources of evidence research-based and in alignment with the state-required sources of evidence (e.g. Professional Growth Planning and Self-Reflection, Observation, Student Voice, and Student Growth Percentiles/ Student Growth)? 📅 Does your district have a plan to support evaluators' selection of sources of evidence in place, with sufficient capacity directed toward the effort? Does sufficient capacity to oversee evaluators' selection of sources of evidence exist? 💰 Are evaluators equipped with the necessary resources (e.g. guidance documents, exemplar plans) to make 	<p>mandatory sources of evidence.</p> <ul style="list-style-type: none"> • Sources of evidence selected by evaluators are inconsistent with state expectations for sources of evidence. • District guidance and partnership with schools on selection of sources of evidence non-existent or limited; district does not access state-provided resources. • District provides limited or no guidance on the selection of sources of evidence to evaluators. • District has not set aside adequate resources or capacity towards supporting evaluators' selection of sources of evidence. • Evaluators' selection of sources of evidence does not meet minimum state-requirements; summative ratings are based on too few or too many sources of evidence without attention to what constitutes a high-quality source of evidence. 	<p>but not limited to, the minimum of Professional Growth Planning and Self-Reflection, Observation, Student Voice, and Student Growth Percentiles/ Student Growth components.</p> <ul style="list-style-type: none"> • Sources of evidence are consistent with state expectations for sources of evidence. • District provides appropriate guidance and support to evaluators when determining sources of evidence and alignment, either via its own channels or through state-provided resources. • District has prioritized resources and capacity toward supporting evaluators' selection of sources of evidence, with a balance of support and challenge in place to ensure evaluators select an appropriate set of sources of evidence. • Evaluators exhibit sound professional 	<p>teacher perspective – neither too few nor too many sources of evidence are included (as revealed via a randomized audit and review of summative evaluations).</p> <ul style="list-style-type: none"> • Guidance documents to support evaluators' selection of sources of evidence have been developed and disseminated. • Feedback from field positive regarding district support and resources is positive; evaluators report receiving clear expectations and support where requested.

Category	Questions to Consider	Weak (1)	Strong (4)	Evidence to look for
	informed decisions about the selection of sources of evidence to inform summative ratings?		judgment by selecting the right balance of multiple sources of evidence to inform their summative ratings; evaluators demonstrate intentionality and attention to what constitutes a high-quality metric when selecting sources of evidence.	
3. Balanced approach– Do evaluators take a balanced approach to the selection of sources of evidence that align with and support the summative ratings in each of the domains of the Framework for Effective Teaching?				
	<ul style="list-style-type: none"> ☐ Has the district developed and/or adopted a clear protocol to help inform evaluators' consideration of sources of evidence (beyond those required) and arrival at a summative rating? • Are evaluators selecting the appropriate number or mix of sources of evidence to inform their summative rating? ☐ Has a clear and well-supported process been established to support local 	<ul style="list-style-type: none"> • There is not a clearly established set of decision rules or guidance from the district to support evaluators' decisions about inclusion and values placed on various sources of evidence. • Evaluators rely on too few or too many sources of evidence/data point, with little evidence that they are considering relative value of each source of evidence and its relationship to a teacher's overall effectiveness. • No clear process in place to support selection of sources of evidence and and/or lack of district 	<ul style="list-style-type: none"> • There is a clear and consistently applied set of decision rules for reviewing and weighing sources of evidence in order to arrive at a summative rating. • Evaluators select a manageable number of sources of evidence and determine a teacher's summative rating based on the relative value of each piece of evidence on a teacher's overall effectiveness. • Selected sources of evidence include following required sources: Professional Growth Planning and 	<ul style="list-style-type: none"> • Student Growth Percentiles (SGPs) are determined using the state's pre-defined cut scores; student growth goals are set in alignment with district-determined rubric. • District provides training and clear expectations about what quality evidence looks like and how it can be effectively aligned with the Framework for Teaching in order to promote an evidence-based summative rating correlated with improved student achievement.

Category	Questions to Consider	Weak (1)	Strong (4)	Evidence to look for
	<p>determinations for sources of evidence?</p> <ul style="list-style-type: none"> • Are evaluators giving appropriate consideration to student growth when determining a teacher's overall performance rating? • Do evaluators demonstrate a nuanced understanding of what "high-quality" and "valid" source of evidence look like? 	<p>resources to support local efforts.</p> <ul style="list-style-type: none"> • Minimal value is placed on student growth when making a determination of a teacher's overall performance category; other, more qualitative sources of evidence are given significantly more consideration. • There is little or no consistency among evaluators' selection and application of sources of evidence; there is an over- or under-reliance on certain sources of evidence at the expense of other valid indicators. 	<p>Self Reflection, Observation, Student Voice, and Student Growth Percentiles and/or Student Growth Goals.</p> <ul style="list-style-type: none"> • Clear process established and utilized to support selection of sources of evidence, with appropriate support from district staff. • Student growth is given significant consideration when determining a teacher's overall performance category. • There is a high degree of consistency among evaluators' selection and application of sources of evidence; sources of evidence are selected with attention to balance of quantity (sufficient number of distinct sources of evidence selected) and quality. 	<ul style="list-style-type: none"> • Clear process established and in use to support local decision-making around weights. • Evaluators are aligned around expectations for the selection and application of sources of evidence (e.g. random audit of summative evaluations reveals a high degree of quality and consistency in the type of evidence leveraged and how it is used to inform the overall summative performance category).

Category	Questions to Consider	Weak (1)	Strong (4)	Evidence to look for
Ensure High-Quality Data				
4. Monitor validity – <i>Is there a process in place to ensure that the selected sources of evidence, when taken together, are associated with improved student learning outcomes?</i>				
	<ul style="list-style-type: none"> • Were sources of evidence selected with consideration to their ability to predict teachers' effectiveness in raising student achievement? 📄 Is validation a one-time event or is there a process in place to monitor the predictive strength of the selected sources of evidence on an ongoing basis?ⁱⁱ 😊 Do you have the capacity and/or expertise to evaluate validity and adjust? 👍 Does your monitoring process result in direct feedback and/or support to evaluators on the selection and application of appropriate sources of evidence? 😊 Across the district, do evaluators select sources of evidence thoughtfully, with attention to their relationship to more 	<ul style="list-style-type: none"> • Few, if any, resources devoted to the careful selection of sources of evidence and the development/ implementation of a process to assess validity of sources of evidence. • If a monitoring process is in place at all, it is intended to be a one-time-only check rather than a periodic review. • Findings from monitoring are resigned to the district office rather than used to inform feedback and support for evaluators. • When evidence suggests misalignment between source(s) of evidence and student achievement gains, it is ignored or incompletely addressed. 	<ul style="list-style-type: none"> • There is well-defined process and adequate resources allocated toward assessing the validity of sources of evidence and/or leveraging existing research on the validity of sources of evidence. • Evaluations are monitored on an ongoing basis to periodically assess the strength of sources of evidence (e.g. are you seeing correlations between high ratings on the sources of evidence identified to inform the professional practice rating and those used to inform the student growth rating?) • Comparisons of student achievement gains with teachers' performance are used to inform decisions about adjusting sources of evidence or, where needed, identifying new ones. • When evidence 	<ul style="list-style-type: none"> • High degree of correlation between student achievement and identified sources of evidence (e.g. those teachers who receive better observation scores and student survey results demonstrate better achievement outcomes for their students than peers with lower scores on the same measures). • Resources are in place to continuously monitor the selection and use of sources of evidence at the school-level. • Evaluators adjust their selection process and professional judgment in accordance to feedback from the district. • Increasingly, the district observes a stronger correlation between teachers' ratings on professional practice, student growth, and student achievement.

Category	Questions to Consider	Weak (1)	Strong (4)	Evidence to look for
	effective teaching and improved levels of student achievement?		suggests misalignment between source(s) of evidence and student achievement gains, there is a process in place and consistently applied to address inconsistencies.	
5. Ensure reliability – Are you confident that sources of evidence consistently reflect the quality of teachers’ practices rather than the biases or idiosyncrasies of a particular rater, lesson, or instrument?				
	<p>📄 Are you aware of and employing “best practices” to reliably measure teachers’ practice (e.g. for observations, multiple observations and observers, observations of different length and at different points in a lesson; for student survey measures and tests, consistent data collection processes, appropriate content, and assurance of confidentiality)?</p> <p>📄 Do you collect and review available data to assess reliability of your system’s sources of evidence?</p> <p>👍 Do you collaborate with KDE to discuss reliability in the field</p>	<ul style="list-style-type: none"> Teachers are observed solely by their supervisor; peer observation component is not employed. Summative evaluation rating is based on single observation rather than multiple lessonsⁱⁱⁱ. Observations are consistently conducted at the same time of day, preventing opportunity to observe different aspects of teaching scored on the Framework for Teaching^{iv}. There is no protocol in place for observers to compare ratings across system. There is little to no training provided to evaluators and/or training is of poor quality; training does not 	<ul style="list-style-type: none"> Each certified teacher receives supervisor and peer observations, providing more than one source of feedback on practice^v. Summative evaluation rating is based on input from multiple observations, ideally conducted at different points in a given lesson. Observations consist of at least one full lesson and several others of varying lengths of time. Evaluators are given the opportunity to compare observation notes and ratings. High-quality training provided consistently to all evaluators; training includes the 	<ul style="list-style-type: none"> Comparisons of ratings across evaluators and lessons yield similar results for a single teacher. Scores on the Framework for Teaching are positively correlated with a teacher’s student achievement gains^{viii} District-developed protocol calls for more than one observer per teacher observation There is a balance between frequency and length of observation. Evaluators feel confident on their ability to rate accurately and reliably, as gathered via feedback loop(s). Evaluators can

Category	Questions to Consider	Weak (1)	Strong (4)	Evidence to look for
	<p>and share feedback and experiences?</p> <p>☐ Is there a robust training procedure in place for evaluators, and is it implemented consistently and with quality?</p> <ul style="list-style-type: none"> Do the sources of evidence at evaluators' disposal result in the identification of effective teaching in a fair and reliable fashion? 	<p>include evidence-based practices such as comparing ratings on video-taped lessons.</p> <ul style="list-style-type: none"> Confidentiality of results from student voice and educator surveys cannot be assured. 	<p>best practice of using videotaped lessons to ensure evaluators' judgment is consistent.</p> <ul style="list-style-type: none"> All evaluators meet minimum requirements for certification and re-certification. There is a process in place to track and monitor evaluator certification and re-certification that includes a scaffolded approach to supporting evaluators who require additional assistance. There is a transparent process to ensure student/teacher confidentiality; stakeholders trust the survey instruments in place^{vi} District-selected observation model (e.g. "Progressive", "Traditional", or "District-Determined") is implemented consistent with KDE-provided guidelines. If district selects "District-Determined" option, the plan 	<p>consistently distinguish between four levels of practice within each domain.</p> <ul style="list-style-type: none"> Distribution of summative ratings differentiates between higher and lower performing teachers.

Category	Questions to Consider	Weak (1)	Strong (4)	Evidence to look for
			reflects best practices such as multiple observations, use of multiple peer observations, and observations of varying length ^{vii} and frequency depending on a teacher's status and prior effectiveness ratings.	
6. Assure accuracy – To what extent does your evaluation process accurately measure effective teaching?				
	<ul style="list-style-type: none"> 📋 Is there a rigorous training protocol in place that requires evaluators to demonstrate their ability to apply the Framework for Teaching prior to formally rating teachers' practice?^{ix} 📋 Do you have process in place to monitor results of evaluations to ensure they produce accurate ratings? 📋 Is there a data verification process in place so that student test scores and survey responses are credited to the right teacher? 😊 Are evaluators consistently able to differentiate 	<ul style="list-style-type: none"> • Principals and peer evaluators are allowed to rate teachers' practice with no certification process or one that is not rigorous. • If there is a certification process, it consists mostly of training about the observation instrument, with few or no requirements for demonstrating the ability to use it. • Evaluators must rely only on the text of the observation instrument and their own instincts to give ratings. • Practices for crediting student test scores and survey results to teachers are inconsistent across the district. • There is no formal process in place to 	<ul style="list-style-type: none"> • Every person responsible for observation is certified to use the selected instrument before they are allowed to rate teachers' practice^x • The certification process is rigorous, requiring every candidate to demonstrate proficiency at observing and differentiating teacher performance against benchmarks provided by experts. • Assessment of rater accuracy is ongoing, even after certification has taken place; opportunities for re-certification over time are required. 	<ul style="list-style-type: none"> • There is rigorous training for every identified observer that emphasizes understanding of the observation instrument, observation skills, and accurate scoring using the instrument^{xiv} • This training^{xv} includes: <ul style="list-style-type: none"> ○ Discussion of the instrument, its competencies, and its performance levels ○ Video examples of teaching for each competency at each performance level ○ Practice scoring videos, with feedback from trainers

Category	Questions to Consider	Weak (1)	Strong (4)	Evidence to look for
	<p>performance across all competencies within the Framework for Teaching?</p>	<p>monitor results of evaluations to ensure they are producing accurate ratings.</p>	<ul style="list-style-type: none"> • There is a clear set of rules for crediting student test scores and survey results to teachers, and it is applied consistently across the system; teachers are given a chance to review and correct errors in student rosters^{xi} • Formal monitoring process implemented routinely to ensure accuracy of evaluation system (e.g. "spot checks" on observers, correlating student achievement data with observation ratings) • When analyses reveal unusual patterns, there is a process in place to investigate and "re-calibrate" as needed^{xii} • District has invested in data system (or leveraged state system) and human capital necessary to monitor accuracy on at least an annual or quarterly basis^{xiii} 	<ul style="list-style-type: none"> ○ Techniques for minimizing rater bias • Raters are required to rate a number of videos pre-scored by experts^{xvi} and achieve a minimum level of agreement with the expert scores in order to be certified • After certification, rater accuracy is maintained through periodic "calibration" against pre-scored videos (Example: Memphis City Schools)^{xvii} • The system uses a roster verification process that includes multiple sources (e.g., administrators and teachers) to credit student test scores and survey results to teachers

Category	Questions to Consider	Weak (1)	Strong (4)	Evidence to look for
Invest in Improvement				
7. Make meaningful distinctions – To what degree does your system differentiate among educators?				
	<ul style="list-style-type: none"> What is your distribution “spread” of teachers across performance categories? Are teachers’ effectiveness ratings distributed equally across performance categories or clustered around the center of the distribution? Is your system, as implemented, resulting in differentiation among educators (e.g. not all educators are rated “accomplished” or “exemplary”)? 	<ul style="list-style-type: none"> Selected sources of evidence yield little or no differentiation among teachers’ effectiveness Teacher performance across the district is clustered in high performing categories, suggesting across-the-board effectiveness that is misaligned with student achievement results Teacher performance tends to be “bunched at the center of the distribution”^{xviii} Teacher performance is “artificially” equally distributed across quadrants 	<ul style="list-style-type: none"> System allows for a fair amount of differentiation among educators’ performance; sources of evidence allow for distinguishing between more- and less-effective teachers. Teacher effectiveness distribution is broadly aligned with student achievement results. 	<ul style="list-style-type: none"> Distribution spread Analysis between educator ratings and student achievement results for the district as a whole, by school, and by evaluator
8. Prioritize support and feedback – Are results used to provide feedback and support to improve educators’ practice?				
a. Data use ¹	<ul style="list-style-type: none"> 👍 Are data used for continuous improvement purposes rather than solely for high-stakes and/or administrative decisions? Are both quantitative and qualitative information about a 	<ul style="list-style-type: none"> Data are being used primarily to make high-stakes decisions Evaluators rely on too narrow a set of data points during feedback conversations (e.g. qualitative coaching based solely on observation). 	<ul style="list-style-type: none"> District and school leaders use data to understand patterns and needs for additional teacher training and support. Evaluators integrate multiple data points, both quantitative and qualitative, into their 	<ul style="list-style-type: none"> There is a shared responsibility for improving teaching Minutes from PLCs and/or district observation of PLCs reveals that data are used as centerpiece of feedback conversation between evaluator

¹ Self-assessment ratings for 8a may not be possible until post-evaluation data are made available. If that is the case, simply skip this segment and note accordingly.

Category	Questions to Consider	Weak (1)	Strong (4)	Evidence to look for
	<p>teacher's overall skillset being leveraged to provide comprehensive feedback?^{xix}</p> <p>👍 When information is collected that reveals a pattern of ineffective teaching practice, are those data used to help those teachers improve practice?</p> <p>📄 How are you using the results to inform type, frequency, and targeting of support to teachers?</p>	<ul style="list-style-type: none"> • School and district leaders have not used data to identify system-wide strengths/weaknesses in teaching practice • Little work is done to understand the link between PD and its utility in the classroom • Little or no investment in district and school capacity and/or routines to use data to inform decisions and as the basis for feedback conversations • Results are simply words on paper and not used to target improvement efforts or form the basis of a meaningful, high-quality feedback conversation. • Formative data (e.g. from benchmark/interim assessments) are not leveraged to inform mid-course corrections throughout the school year. • Evaluators do not provide meaningful, "just in time" feedback; feedback fails to pinpoint specific opportunities for growth^{xx} 	<p>feedback conversations (e.g. data from student surveys, scores from multiple observations that reveal a trend in a specific dimension, a qualitative coaching conversation).</p> <ul style="list-style-type: none"> • Leadership discussions rely heavily on data and decisions are made based on that data. • Effective routines are in place at the district and school level (e.g. PLCs) to provide a venue for rich, data-driven discussions • Adequate capacity and training provided to district and school leaders to serve as strong data "stewards" and model behavior and skills. • District and/or schools use findings to target resources to support and improve practice within specific dimensions of the Framework for Teaching^{xxi} • Targeted professional learning opportunities reflect state-adopted 	<p>and teacher, as a "flashlight" to guide improvement rather than a punitive "hammer."</p> <ul style="list-style-type: none"> • Feedback from educators (via TELL and other surveys) reveals feedback conversations to yield a high degree of value and result in good recommendations for how to improve practice. • Professional development offerings (type of PD, to whom it is offered) are determined based on where teachers fall along effectiveness distribution; teachers who need the most support receive it. • Audit of written feedback to teachers includes specific, actionable recommendations that can be implemented in the classroom immediately.

Category	Questions to Consider	Weak (1)	Strong (4)	Evidence to look for
			professional learning standards. <ul style="list-style-type: none"> Formative data collected and used on an ongoing basis to inform instructional decision-making Evaluators provide meaningful, “just in time” feedback; feedback is specific and actionable. 	
b. Type and mechanisms for feedback	<ul style="list-style-type: none"> Do you have feedback loops in place to provide your system with timely and relevant information about implementation status? <ul style="list-style-type: none"> How do you know, at any given moment, whether you are on-track? What procedures/ routines do you have in place to facilitate implementation efforts? Is there an accountability mechanism in place not only for compliance (e.g. conducting the appropriate number of classroom visits), but for delivering timely 	<ul style="list-style-type: none"> Little information (beyond an immediate evaluation) is gathered on the utility of professional learning opportunities and resources Decisions are being driven by anecdotal information Peer observer role not employed, or not employed with fidelity. Little or no accountability for delivering useful feedback that helps teachers improve practice. 	<ul style="list-style-type: none"> Leaders have created a set of mechanisms, including additional observations, surveys, and analyses of formative assessments, to understand how instructional practice is changing Qualitative information about professional learning resources are gathered and shared routinely with district leaders Peer observers used regularly to provide timely formative feedback on practice Mechanisms exist that link interim student progress with instructional supports Accountability 	<ul style="list-style-type: none"> Feedback loops to assess the quality of implementation of PGES and its various components are in place; results used to inform current and future implementation efforts, professional development, and support to practitioners. Formal protocol and clear expectations for use of peer observation. Documentation on peer reviewer certification Teacher reports on quality and timeliness of feedback

Category	Questions to Consider	Weak (1)	Strong (4)	Evidence to look for
	and meaningful feedback? ^{xxii}		emphasizes quality of feedback rather than basic compliance.	
9. Use data for decisions at all levels – <i>Does strategic use of data enable your system to better support teachers' improvement needs?</i>				
	<p>☐ How are data about teachers' effectiveness across domains being used to inform the spectrum of human capital decisions (e.g. hiring, placement, retention/promotion, professional development)?</p> <p>\$ Is there alignment between professional development resources and areas of teaching that classroom observation measures indicate need improvement most?</p>	<ul style="list-style-type: none"> • Little or no investment made in training and supporting stakeholders' awareness, use, and analysis of data on teacher effectiveness • There are few or no routines in place to discuss results and make data-driven decisions (e.g. PLCs at school and/or district level). • Professional development opportunities are made in a vacuum, without attention to evaluation results across a school and/or district, or attention to areas of teaching that summative results indicate need the most improvement. • Spending decisions are uninformed by data, and are made on an ad hoc basis. 	<ul style="list-style-type: none"> • Considerable investment is made in human capital, training, and support to facilitate strategic data use across at all levels. • Well-implemented routines are in place with wide participation at all levels to analyze and make decisions based on data. • Data are collected and reviewed on an ongoing basis to determine type, frequency, and impact of professional development. • Decisions about how to spend limited resources are informed by evaluation results. • Capacity to routinely analyze evaluation results and identify common areas of need across schools/district exists. 	<ul style="list-style-type: none"> • Stakeholders at all levels demonstrate awareness of available data and are skilled in analysis and application to human capital decisions. • There is a shift away from "one size fits all" professional development to targeted opportunities based on identified needs. • There is a comprehensive survey taken of the type, frequency, and characteristics of feedback conversations between evaluators and teachers, as well as among district staff; discussions should be rooted in data and evidenced-based.

ⁱ HOLD FOR CITATION

ⁱⁱ “Feedback for Better Teaching,” Page 6.

ⁱⁱⁱ “Culminating Findings from the MET Project’s Three-Year Study”: Figure 5, Page 18

^{iv} “Ensuring fair and Reliable Measures of Effective Teaching,” Page 19.

^v “The Reliability of Classroom Observations by School Personnel,” Table 10, Page 29

^{vi} “Culminating Findings from the MET Project’s Three-Year Study,” Page 20.

^{vii} “The Reliability of Classroom Observations by School Personnel,” Page 13

^{viii} “The Reliability of Classroom Observations by School Personnel,” Page 4

^{ix} “Feedback for Better Teaching,” Page 6.

^x “Gathering Feedback for Teaching” Policy and Practice Brief: Figure 4, “Ensuring Accuracy of Observers”

^{xi} “Culminating Findings from the MET Project’s Three-Year Study,” Page 13

^{xii} “Ensuring Accurate Feedback from Observations: Perspectives on Practice,” Page 27

^{xiii} “Ensuring Accurate Feedback from Observations: Perspectives on Practice,” Page 28

^{xiv} “Foundations of Observation,” Page 14

^{xv} “Gathering Feedback for Teaching” policy and practice brief, Page 14

^{xvi} For more on how such videos can be created (or identified from outside experts), see “What It Looks Like: Master Coding Videos for Observer Training and Assessment,” particularly Figure 1.

^{xvii} Ibid., Page 16

^{xviii} “Feedback for Better Teaching: Nine Principles for Using Measures for Effective Teaching,” Page 7

^{xix} “Ensuring Accurate Feedback from Observations: Perspectives on Practice,” Page 9

^{xx} “Ensuring Accurate Feedback from Observations: Perspectives on Practice,” Page 11

^{xxi} “Ensuring Accurate Feedback from Observations: Perspectives on Practice,” Page 9

^{xxii} “Fixing Classroom Observations,” TNTP, Page 3