

Reimagining Instructional Coaching

Developing Observation Tools
to Support Instructional Coaching
in Pre-K Classrooms



Introduction

3

Criteria and Thresholds of Observation Tools to Support Improvements in Classroom Practice

Goal 1: Content	10
Goal 2: Psychometrics	15
Goal 3: User Experience	18
Goal 4: Usefulness	21
Goal 5: Scalability	24

Appendices

28



Introduction

Decades of evidence demonstrates that enrollment in high-quality, center-based preschool programs can boost the development of children's early skills and provide a solid foundation for their later school success.¹ A high-quality early childhood classroom provides children from all communities and families with a safe and nurturing environment that facilitates equitable opportunities to develop their full range of physical, social, emotional, and academic capabilities.² Unfortunately, not all preschool programs provide sufficient educational quality to yield these benefits.³

Instructional coaching, informed by observation tools that measure teachers' practices, has been effective in improving teaching quality in early learning programs.⁴ Observation tools represent a critical component of the coaching process as they enable instructional coaches and teachers to identify specific strengths and challenges in teachers' instructional practices and allow teachers to receive regular feedback on their progress toward improving these practices.⁵ However, existing measurement tools limit teachers' abilities to implement this type of instructional coaching at scale.⁶

Introduction *(continued)*

The observation tools most commonly used to support instruction do not fully attend to culturally responsive practices and biases related to the demographic, cultural, and linguistic backgrounds of teachers and children. Many instructional coaches rely on observation tools designed for formal assessments that require extensive training and are expensive to implement. This limits the timeliness, frequency, and specificity of feedback needed to work on changes to behaviors associated with various teaching practices.

This document helps to address that challenge by detailing a Target Product Profile (TPP) that describes the ideal features and capabilities of observation tools to support instructional coaching with prekindergarten (pre-K) teachers. This TPP prioritizes the experiences of pre-K teachers, families, and children whose perspectives are less often centered in the design, creation, and validation of measurement tools—such as Black and Latine children,⁷ children who are multilingual,⁸ children with disabilities, and children from low-income backgrounds.

This TPP has two primary purposes. The first is to spur investments, improvements, and innovations in existing and future observation tools for instructional coaches that measure important teaching practices and provide useful feedback of ways to improve classroom quality (actions and/or resources). The second purpose is to support collaborations between early learning decision-makers, assessment suppliers, researchers, teachers,⁹ and families with children in early education programs to develop tools for instructional coaches that promote continuous improvement in pre-K classrooms.

The Bill and Melinda Gates Foundation supported the creation of this TPP to advance the development of accurate, equitable, useful, and innovative tools to support coaching focused on the quality of center-based pre-K classrooms that serve 3- to 5-year-old children in the United States. This TPP was informed by, but is distinct from, a separate TPP that focuses on child assessments to track children's growth in a range of important developmental domains.¹⁰ The teams that produced the two TPPs followed similar approaches to gathering information and determining the respective goals, subgoals, criteria, and thresholds.

Vision for New Observation Tools to Support Instructional Coaching

Well-designed and properly implemented classroom observation tools can support teachers in providing high-quality interactions and learning environments, enabling children to reach their full potential. To do this, tools must be useful, scalable, and generate actionable, timely, accurate, and reliable information for teachers and instructional coaches. Assessment developers must gain a deeper understanding of how various stakeholders interact with and are impacted by the tool. These stakeholders include actors and beneficiaries.

ACTORS

Those who are actively involved with the use of the coaching tool.

- **Classroom teachers.** Lead and assistant teachers who interact with children daily. The outputs of the observation tool will provide information on the strengths and growth opportunities in their teaching.
- **Instructional coaches.** Coaches, program administrators, principals, professional development specialists, and peer teachers who observe teachers and provide guidance on teaching practices. Paired with classroom teachers, these coaches have the most direct relationship to every step of the process of using the tool.
- **Program administrators/Policy makers.** Leaders of preschool programs and policy makers who determine program-wide priorities, including professional development. The outputs of the observation tool will affect decision-making at the program level.

BENEFICIARIES

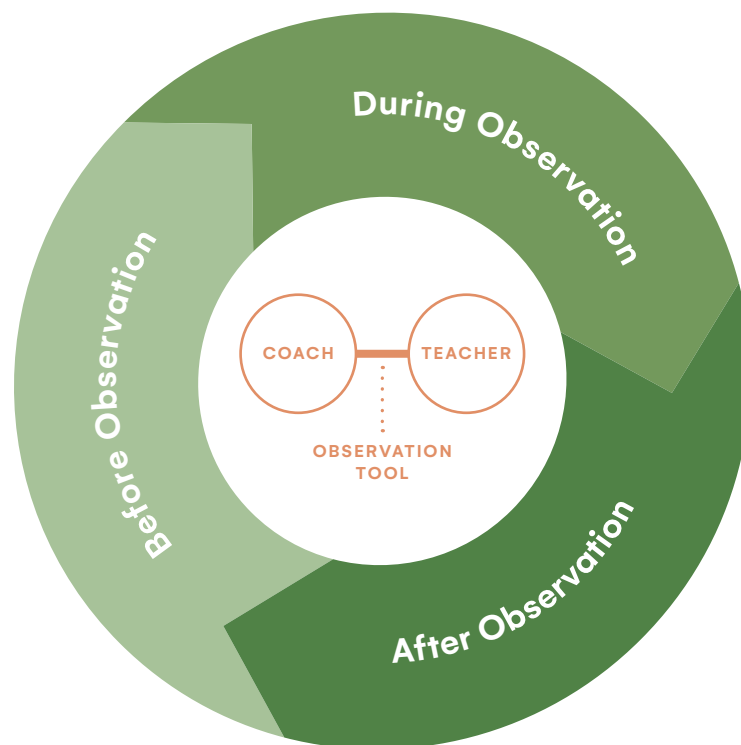
Those who benefit from the use of the coaching tool.

- **Classroom teachers.** Teachers who receive targeted feedback and guidance on areas for improvement are better equipped to enhance their teaching skills and increase their confidence in their abilities. This, in turn, can lead to greater job satisfaction and overall well-being, making them more likely to remain in the teaching profession.
- **Children.** Children are present in the classroom environment during observation and are therefore affected both directly and indirectly by the tool's use. As teachers grow in their practice, children directly benefit from higher quality interactions and learning environments.



Instructional Coaching Journey

An effective instructional coaching journey connects classroom teachers, instructional coaches, and program administrators, with a tool supporting their experience before, during, and after observation. The process can also fold into a cohesive system of classroom observations, which include formal observations such as the Classroom Assessment Scoring System (CLASS) and Early Childhood Environment Rating Scale (ECERS).



Before Observation



Instructional coaches and teachers meet to:

- Discuss teacher self-assessment of classroom strengths and challenges
- Share classroom content
- Determine goals and time for classroom observation

Tool:

- Supports shared decision-making on focus of observation from a range of evidence-based teacher practices
- Generates an observation guide to focus on teacher strengths and areas of improvement

During Observation



Instructional coaches:

- Collect information aligned to the predetermined focus of the observation

Teachers:

- Implement lessons as planned

Tool:

- Provides guidance on what to “look for” during the observation
- Guides instructional coach to observe how individual children experience the classroom (e.g., Black, Latine, and multilingual children, and children with disabilities)

After Observation



Instructional coaches and teachers meet to:

- Reflect on observation findings
- Provide actionable feedback
- Refine goals for continued improvement

Tool:

- Supports communication of observation data focused on strengths and opportunities for improvement
- Provides resources aligned with the observation findings

Administrators:



- Reflect on whether coaching areas of focus are showing improvement in formal observations
- Determine areas for program-level professional development and coaching

Tool:

- Generates reports that identify strengths and opportunities for program-level improvement

What is a Target Product Profile?

A Target Product Profile (TPP) outlines design parameters that reflect the priorities identified by end users, as well as research and best practices where they exist. These parameters support the identification of needs where further innovation, pilot-testing, and iteration on observation tools in collaboration with a broad range of stakeholders are needed. They prioritize the experiences, strengths, and needs of pre-K teachers, families, and children whose perspectives are less often elevated in the early design, creation, and validation of measurement and assessment tools—in particular, Black and Latine children, children who speak both Spanish and English, and children experiencing poverty.

Ways to Use This Target Product Profile

This document will help assessment developers, funders, and scholars develop the next generation of observation tools for instructional coaches by supporting the following activities:

- **Development of item content or solutions.** To identify item content that will assist in achieving the outlined goals.
- **Gap analysis.** To identify places where existing observation tools do not address the goals.
- **Prioritization.** To identify focus areas and features in the context of the development of new tools.
- **Equitable design.** To develop features, address gaps, and prioritize culturally responsive solutions and approaches.
- **Solution ideation.** To drive sessions in which product teams generate measurement solutions via possible methods such as co-design sessions, workshops, and team exercises.

For institutional decision-makers, this document identifies important attributes for evaluating existing and selecting new instructional observation tools that meet or exceed the thresholds for performance.

How This TPP Should Not Be Used

- This TPP should not be viewed as static or overly prescriptive; it does not articulate the process by which observation tools for instructional coaches should be revised or created. Instead, the TPP provides the structure for setting a more open, creative, and iterative approach in motion, an approach that seeks to address gaps in research and promote more equitable processes for ideation and testing in the field.
- New and revised tools for instructional coaches should not be viewed as solutions to long-standing educational inequities; rather, these products should be seen as important levers for improving learning and developmental outcomes for all pre-K children.
- The tools developed using the TPP should not be used for high stakes monitoring and evaluation of teachers. The TPP aims to support the development of tools that will help instructional coaches and teachers focus on improving impactful pre-K experiences for all young children.

Development Methodology

This TPP was developed by SRI Education, a division of SRI International, and Substantial, with funding support from the Bill and Melinda Gates Foundation's Early Learning Team, and with contributions and engagement of teachers, instructional coaches, families, program administrators, state and local pre-K systems administrators, academic experts, and innovators and entrepreneurs in the pre-K field.

Data sources included interviews, focus groups, ideation workshops, and a review of prior research on early childhood classroom observation tools. Participants in these activities included families with young children and teachers working in Head Start, community-based childcare, and other publicly funded pre-K settings.¹¹ Other participants included academic and pre-K operations experts, including assessment researchers, developmental psychologists, nonprofit leaders, heads of agencies serving young children, experts on multilingual children (particularly those who speak Spanish), as well as an array of pre-K decision-makers, advocates, and thought leaders from more than 20 geographically diverse states and localities. Across all participants, we prioritized the perspectives of Black families, Latine families, families of children who are multilingual learners, and families who experience poverty. These groups warrant specific attention because their historic marginalization denies children access

to positive learning experiences that affect their school outcomes. It should be noted there is substantial heterogeneity within these groups, and their experiences do not reflect the experiences of all historically marginalized groups of children.¹² However, they represent a starting point for ensuring the development of more equitable and inclusive tools that meet the needs of a broad range of families and children being served in publicly funded pre-K programs today.

The design parameters outlined in this TPP reflect the suggestions of research and best practices and highlight areas for further innovation, pilot-testing, and iteration on future observation tools in collaboration with teachers, program administrators, policy makers, and assessment developers. We nevertheless see this as a living document to be updated to align with insights from research and to reflect evolving policy and family priorities.

Attributes of the Pre-K Classroom Observation Tool Target Product Profile (TPP)

The following attributes are highlighted as important thematic areas of the TPP to address in developing observation tools for instructional coaches focused on classroom quality.

1. **Content:** Pertains to characteristics of teaching, learning environments, and interactions that should be measured to support high-quality early childhood classrooms.
2. **Psychometrics:** Has sufficient evidence to support the use of the observation tool for its intended purpose, its ability to capture meaningful differences in classrooms, its sensitivity to changes in teacher practice, and the relationship of observation-produced scores to critical child outcomes. In addition, the extent to which the tool generates data that can be used equitably with minimal bias across classrooms that serve children from various racial, ethnic, and income identities, who speak languages other than English, and live in different geographic areas within the United States, and across teachers of various racial, ethnic, and geographic backgrounds.
3. **User Experience:** Reflects the optimal experience of important stakeholders (pre-K teachers, instructional coaches, program administrators, children, and their families) using or benefitting from the observation tool and the resulting data and its use to improve classroom quality.
4. **Usefulness:** Supports meaningful use of the data and findings for important stakeholders, including teachers, instructional coaches, and program administrators. Data for teachers and programs will inform classroom-level improvements in children's experiences in the pre-K classroom through feedback, related coaching and professional development, and targeted resources.
5. **Scalability:** Enables easy training and implementation, affordable expansion of the tool for instructional coaching, updates to data systems without losing data, and ability to link coaching data with local data and professional development systems.

Within each thematic area, we articulate a set of goals and subgoals, criteria that list the general design parameters for the subgoal identified, and thresholds that provide indicators or suggested practices to assess the degree to which subgoals are successfully met.

Content

Observation tool measures the quality of 3- to 5-year-old children's experiences in pre-K classrooms in equitable and culturally responsive ways, including the experiences of children who are Black, Latine, and multilingual, and who have disabilities.

Pre-K classrooms provide young children with opportunities to develop the social and academic skills that can set them on a trajectory for lifelong success. Doing this well is a complex task. Pre-K teachers serve children from a diverse array of backgrounds and deploy a broad set of knowledge and skills while adapting to the ever-changing strengths and needs of their students. Pre-K program administrators, instructional coaches, and teachers need clear guidance on how to foster high-quality experiences for children. Developers of classroom observation tools would also benefit from learning from pre-K teachers on what their conceptions of classroom quality entail. Observation tools of pre-K classroom quality must capture the adult—child interactions and environmental features that predict growth on important child skills and outcomes for children of diverse racial, linguistic, and family income backgrounds. Equity is central when determining what is being measured and how it is being measured. Many current conceptions of high-quality pre-K have been dominated by a white middle-class worldview. Children's cultural backgrounds influence the ways they demonstrate interests, imitate others, or engage in play situations; drawing on children's skills and experiences will help them feel known, safe, and more interested in learning.¹³ Moving forward, definitions of classroom quality must reflect the cultural values, practices, and strengths of Black, Latine, and multilingual children and families. Observation tools must also proactively address issues of racism and ableism. The subgoals listed under Goal 1 are drawn in part from the principles outlined in *Measuring the quality of early learning environments: A guide to evaluating ideal learning environments for young children*.¹⁴

Goal 1: Content

SUBGOAL 1.1	The observation tool measures the extent to which teachers provide equitable and inclusive instruction.	
ID	CRITERIA	TARGET THRESHOLDS
1.1.1	Teachers' practices are reflective of the cultural perspectives and practices of Black, Latine, and multilingual families.	— Developers use frameworks/best practices from community-based participatory methods to collect and use input from a range of end users including parents and caregivers, particularly those from marginalized backgrounds, as well as educators and program leaders in the development of the tool to ensure it reflects the strengths and experiences of teachers and children with varied demographic, linguistic, or socioeconomic backgrounds. — The observation tool includes elements of teacher practices and classroom features considered indicative of a high-quality classroom environment by parents/guardians/caregivers, early educators, and Advisory Group members with expertise in equity-informed and culturally responsive literature and research methods with Black, Latine, and multilingual families.
1.1.2	Teachers' practices account for children's culture, strengths, and identity.	— The observation tool captures information on whether: • Teachers vary their approaches to working with children to best capitalize on their funds of knowledge (e.g., individual, linguistic, and cultural strengths). • Teachers make active connections between the children's community, language, and culture and early learning content and pedagogy. • Teachers provide classroom materials (e.g., books, artwork, dolls) and activities (e.g., music, cooking, thematic play) in the program that draw from and incorporate a diversity of cultures and languages and include positive representations of persons with disabilities.
1.1.3	Children with different characteristics have equitable opportunities and experiences.	— The observation tool captures the extent to which there are differences in the opportunities and experiences of children based on child characteristics (including Black and Latine children, multilingual children, and children with disabilities). This may include the number and type of interactions classroom staff have with different groups of children (e.g., ask questions, invite to participate in an activity or to answer a question).

Goal 1: Content

1.1.4	Teachers engage families as partners in their children's learning at school and home.	— The observation tool measures the use of family engagement practices and policies.¹⁵ These may include but are not limited to: • Teachers invite families to participate in decision-making and goal-setting for their children. • Teachers engage families in two-way communication. • Teachers provide learning activities for the home and in the community. • Teachers invite parents to lead and support planning and implementation of activities for the class and other families. • Teachers learn about family needs and barriers. • Teachers communicate positive messages about children's families and home practices.
1.1.5	Teachers provide evidence-based supports for multilingual children.	— The observation tool measures the use of evidence-based practices¹⁶ to support the development of multilingual children in pre-K. These may include but are not limited to: • Teachers incorporate children's home languages in instruction and classroom materials. • Teachers provide multilingual children with language scaffolding (e.g., anchor words, sentence stems). • Teachers incorporate visuals into instruction to support understanding of vocabulary. • Teachers use layered questioning techniques in which they begin with a question that can be answered with a simple fact or definition and follow up with a series of increasingly complex questions, each one building on the previous answer.
1.1.6	Teachers provide evidence-based supports for children with disabilities.	— The observation tool assesses the use of strength and evidence-based practices¹⁷ to support the development of children with disabilities. These may include but are not limited to: • Teachers ensure all materials (toys, supplies, equipment) are accessible to all children. • Teachers adapt materials and instruction to meet specific needs of children with disabilities. • Teachers provide all children with multiple means of engagement, representation, and expression. • Teachers identify and eliminate barriers that limit children's access to the learning environment.

Goal 1: Content

SUBGOAL 1.2	The observation tool measures the extent to which teachers' use of evidence-based practices support the development of preschool-age children.	
ID	CRITERIA	TARGET THRESHOLDS
1.2.1	Teachers support children's positive relationships with adults, peers, and self.	— The observation tool captures indicators of teacher practice such as: • Teachers interact with children in a warm, positive, and encouraging manner. • Teachers facilitate positive behavior and peer relations. • Teachers model effective social skills and behavior. • Teachers provide supports to respond to children's specific emotional needs. • Teachers understand children's family experiences and respond accordingly. • Teachers and classrooms have routines that support transitions to/from learning opportunities. • Teachers do not use harsh or exclusionary discipline such as being sent out of the classroom or home as a form of punishment. • Teachers can recognize symptoms of trauma and can provide trauma-informed daily routines and practices.
1.2.2	Organization/structures intentionally facilitate children's exploration, independence, and interaction.	— The observation tool captures indicators of teacher practice such as: • Teaching and activities (whole-group, small-group, and free-time stations) are well prepared and engaging. • Teachers recognize, support, and encourage children's playful learning by providing resources and opportunities for children to engage in deep, interactive, and creative play. • Teachers and classrooms have routines that support transitions to/from learning opportunities. • Teachers and classrooms have routines that support children's engagement during activities.
1.2.3	Teachers serve as guides, nurturing presences, and co-constructors of knowledge.	— The observation tool captures indicators of teacher practice such as: • Presence of instruction for specific learning domains: math, literacy, language development, science, physical development, social science, music, movement, and art. • Teachers scaffold new content knowledge. • Teachers guide children as they conduct their own exploration and investigation of ideas and concepts, rather than solely providing didactic instruction. • Teachers engage in practices that support children's positive self-efficacy, perseverance, and sustained focus.

Goal 1: Content

ID	CRITERIA	TARGET THRESHOLDS
1.2.4	Teachers personalize instruction to each child's development and abilities.	— The observation tool captures indicators of teacher practice such as: • Teachers provide children with opportunities to engage in a diverse set of activities that align with their interests and needs. • Teachers provide learning opportunities in a range of formats. • Teachers document children's development and use that information to personalize learning activities with small groups or individual children.

GOAL
2

Psychometrics

Sufficient validity evidence is collected to support the interpretation and use of feedback from the observation to support teacher and/or classroom practices.

The purpose of an instructional coaching tool is to provide feedback that coaches and teachers can use to improve teaching and/or classroom practices. In turn, this should indicate improvements in practice that are reflected with positive child outcomes, either directly or through improvements in formal classroom quality assessments that are associated with positive child outcomes. It is important to ensure the tool measures what is being observed and that the feedback associated with these observations is aligned with the constructs the tool is designed to measure (e.g., evidence of test content validity is collected through typical use). Therefore, evidence related to reliability and validity¹⁸ should be collected to ensure the observation tool is providing accurate feedback that can be used for instructional coaching. In addition to collecting reliability and validity evidence, assessment developers need to demonstrate evidence of the fairness of the assessment and should take explicit efforts to measure and minimize bias. Although we expect to see differences in classrooms implementing different practices, the assessment should minimize bias related to construct-irrelevant differences.

Goal 2: Psychometrics

SUBGOAL 2.1	Sufficient validity evidence is collected related to consistency in understanding of test content and observer responses generated.	
ID	CRITERIA	TARGET THRESHOLDS
2.1.1	Evidence indicates there is consistency with how observers interact with the observation tool and the expectations of that interaction.	— Qualitative analysis of how raters respond to or use indicators demonstrates consistency with the intent of the indicators and/or observation (e.g., cognitive interviews provide evidence that the thought process used by raters is as intended). — Validation study data that include a range of participants demonstrate that with the recommended training, coders reach acceptable kappa scores (greater than 0.61 indicating good reliability¹⁹) with a set of scores provided by experts on a set of videos that include a range of contexts and situations.
2.1.2	Scores and/or feedback are consistent when there are no changes to instructional practices, classroom environment, or other aspects of children's experiences in the pre-K classroom.	— Observation tool produces acceptable test/retest reliability evidence, such as scores that are in the same range and/or similar feedback (i.e., have the same overall interpretation) when there are no changes in the classroom. For example, a study involving a range of participants demonstrates that observers are able to achieve 80% reliability on a set of indicators (module) with their own rating on the same video, based on observations done at different time points.
2.1.3	Scores and/or feedback from the observation tool are not biased toward any racial or ethnic orientation (teachers and/or children).	— Differential item functioning (DIF) analysis on demographic groups of teachers and/or classrooms demonstrates comparability of scores (e.g., less than 5% when context is comparable). This requires that demographic information be collected before conducting this analysis and that there are enough teachers in each demographic group to be comparable. Scores would be compared for a given rater to ensure the tool is not biased. — Information from the observation tool is reviewed at regular intervals to ensure constructs remain reliable and the content is not biased over time across prioritized populations. Observation tool is updated appropriately if needed. — External review by a panel that includes experts in equity in pre-K classrooms determines the indicators are not biased toward any one demographic, linguistic, or socioeconomic group of teachers or children.

Goal 2: Psychometrics

SUBGOAL 2.2	Sufficient validity evidence is provided to demonstrate that the measure is sensitive to variation and changes in teacher practice and relates to child outcomes.	
ID	CRITERIA	TARGET THRESHOLDS
2.2.1	There is evidence that the observation tool captures differing levels of mastery of teacher practices.	— There is variation across teachers in how they perform on the observation tool (e.g., scores follow an approximately normal distribution, teachers with different levels of skills receive different feedback). — A panel of experts in early childhood education come to agreement that the levels of mastery assigned to various levels of performance on indicators are appropriate.
2.2.2	Scores determined at a specific point in time predict (controlling for baseline and factors) child outcomes and/or scores on a formal classroom assessment related to child constrained and/or unconstrained skills.	— Children’s classroom experiences are observed (e.g., learning/instruction, teacher—child and peer-to-peer interactions, instructional materials, classroom climate, classroom organization/structure) and are associated with children’s performance on assessments related to constrained and/or unconstrained skills. Note that one way to demonstrate this relationship is to find a significant relationship between the instructional coaching tool and a formal tool that has been shown to be predictive of child outcomes.
2.2.3	Improvement in teachers’ scores across multiple time points on the observation measure is (positively) associated with changes on child outcomes and/or scores on a formal classroom assessment that is associated with improvements on child outcomes.	— Children’s classroom experiences are observed across time points (e.g., learning/instruction, teacher—child and peer-to-peer interactions, instructional materials, classroom climate, classroom organization/structure), and changes in scores are associated with changes in children’s performance on assessments related to constrained and/or unconstrained skills. Note that one way to demonstrate this relationship is to find a significant relationship between changes in scores on the instructional coaching tool and changes in scores on a formal tool that has been shown to be predictive of child outcomes.
2.2.4	Scores on the tool improve with coaching or other professional development that focuses on the indicators measured by the observation tool.	— Studies indicate that scores and/or feedback on the observational measure increases when teachers enact improved practices in their classrooms. — A literature review provides support that the type of feedback and/or guidance provided from the use of the observation tool is related to improvement in teacher practice.

GOAL

3

User Experience

Observation tool implementation process is unobtrusive, equitable, and supportive.

The introduction of adults who are not in the classroom on a daily basis, such as instructional coaches, can generate anxiety among teachers and reduce the accuracy of any classroom quality observation. Observation tools designed to support continuous quality improvement are most effective when they provide teachers with agency, such as decisions for the focus and timing of the classroom observation session, direct access to their own data, and the opportunity to provide relevant context to instructional coaches. Observation tools must therefore be designed to support teacher agency and shared interest in the continuous improvement process. Further, to provide accurate and useful feedback, instructional coaches need to understand how to apply the observation tool in pre-K classrooms that use different types of curricula and pedagogies and educate children with different backgrounds and home languages. To support instructional coaches, observation tools should be easy to implement efficiently and accurately in a variety of contexts.

Goal 3: User Experience

SUBGOAL 3.1	Classroom-level observation tool maximizes engagement of teachers.	
ID	CRITERIA	TARGET THRESHOLDS
3.1.1	Teachers receive clear information about the purpose and use of the observation for improvement prior to observation.	— Observation process provides teachers with clear information on: • The observation purpose and focus prior to the observation. • The observation content and metrics prior to the observation. • When and in what format they will receive observation feedback.
3.1.2	Teachers have agency regarding the focus of the observation.	— Classroom observations enable the use of discrete subscales or modules that align with areas of teaching practices and interactions on which teachers and coaches choose to focus. — Observation process provides teachers with the opportunity to collaborate with the observer to determine the content or specific aspect of classroom instruction that will be measured.
3.1.3	Teachers have agency regarding the timing of the observations.	— Observation process provides teachers with the opportunity to collaborate with the observer to determine the ideal time for observation. — Teachers receive at least 2 days' notice that an observation will take place.
3.1.4	Teachers can provide contextual information to the observer before or after the observation.	— Observation process provides teachers with a mechanism (e.g., pre or post interview or feedback form) to support the observer's understanding of classroom dynamics during the observation window. This may include providing contextual information about a child, family, or teachers; antecedents to activities and interactions that were witnessed by the observer; and/or the teacher's own observation of aspects of classroom quality.

Goal 3: User Experience

SUBGOAL 3.2		
Observation tool and training support use across a range of classroom contexts.		
ID	CRITERIA	TARGET THRESHOLDS
3.2.1	Observation tool can be flexibly implemented to different language contexts.	— Prior to observation, teachers provide information on the primary language of instruction (e.g., Spanish, Portuguese, American Sign Language). Only observers who are fluent in the language of instruction can observe the classroom. — Observation tool provides guidance to observers regarding how to mitigate bias when using the tool in classrooms in which teachers and children have accents and use dialects.
3.2.2	Observation tool can be flexibly implemented to different pedagogical and curricular contexts.	— Platforms, training materials, and observation tool include information on how the target instructional practices and classroom structures may be enacted through different curricula and teaching methods.
3.2.3	Trainings and observation tool support the use of the tool across classrooms with children and staff with different backgrounds.	— Training materials and rubrics for ratings include culturally representative examples from classrooms with Black and Latine teachers and children and show different ways that quality can be demonstrated in varied cultural contexts.

GOAL

4

Usefulness

Observation tool provides easy-to-understand and actionable feedback.

The purpose of *observing* for quality in early learning classrooms is to *improve* the quality of early learning classrooms. When observation scores are used to directly inform teacher professional development, the goal is to ensure children have access to educational experiences that support their short- and long-term development. To achieve this goal, teachers, coaches, and program directors need information that is clear, relevant, and actionable for their specific needs. Observation tool developers must provide scores and/or feedback to users within a strengths-based framework that supports their understanding and application of this information to improve children's experiences in pre-K classrooms.

Goal 4: Usefulness

SUBGOAL 4.1	Observation tool generates actionable information for teachers and coaches.	
ID	CRITERIA	TARGET THRESHOLDS
4.1.1	Observation tool identifies how findings are related to specific instructional practices of classroom staff.	— Observation tool findings are mapped onto clear actions that teachers can take to improve. Examples of teacher actions may include increasing the use of advanced vocabulary, serve and return interactions, and delivery of specific content. — Observation tool reports for observers and/or professional development specialists provide links to guidance on how they can use observation tool results to support teachers (e.g., questions and actions to consider). — Observation tool should provide guidance to the coach on how to support the teacher—for example, “Based on what you observed today, these are the key action items to discuss with the teacher.” — Teachers have access to resources to help them understand and apply their findings to address their personal and specific instructional goals (e.g., interactive dashboard with which teachers can see their results next to identified thresholds).
4.1.2	Reports are easy for teachers to understand.	— Most teachers (at least 80%) report being able to understand information from reports quickly and efficiently. — Automated reports are at or below eighth grade reading proficiency and do not require advanced training to understand.
4.1.3	Observation tool findings support classroom and program staff to understand and apply the results for improvement over time.	— Report and feedback templates use a strengths-based framework. This may include narratives of observed positive practices that teachers can build on to improve overall quality. — Observation tool permits teachers to track progress toward criterion-referenced standards within and between years. — Findings can be examined together with child and program observations to see how they might inform target areas of growth.
4.1.4	Observation tool findings are available to teachers and coaches in a timely fashion.	— The observation tool enables observers to provide teachers with some feedback immediately upon the conclusion of the observation. — Observation tool enables observers to provide full observation scores and feedback to teachers within 1 week of the observation.

Goal 4: Usefulness

SUBGOAL 4.2		
Observation tool generates actionable information for program staff.		
ID	CRITERIA	TARGET THRESHOLDS
4.2.1	Observation tool identifies opportunities to adjust program policies to support improved teacher practice.	— Observation tool findings identify aspects of quality that can be improved through revisions to program-level policies. Aspects of quality that are features of the program may include physical structure of the classroom, availability of materials, curriculum, and instructional or classroom management practices mandated by the program.
4.2.2	Observation tools produce results that can be used for program-level planning.	— Observation reports could include but are not limited to: • Program-level view of which skills/practices teachers have been coached on. • Aggregate progress over those skills/practices across teachers in the program. • Connections to formal observation scores that also protect teacher confidentiality.

GOAL
5

Scalability

Observation tool can be administered at scale in a variety of classroom contexts. This includes that the observation, associated training, and available reports are convenient to use, are affordable, and can operate well with existing data systems.

There is a pressing need for sound information on the quality of children's early learning experiences. Classroom observation tools must therefore be able to be deployed in a broad set of pre-K programs. Observation tools must be adaptable to various system goals and quality improvement initiatives and be interoperable with other data systems to promote a cohesive system of classroom-level improvement. To support wide adoption and frequent use, observation tool training and administration must be affordable and efficient. The observation tool needs to fit within the existing structures (e.g., schedules) of programs. Online platforms can support training and observation analyses, generate actionable reports, and track progress over time.

Goal 5: Scalability

SUBGOAL 5.1	Observation tool implementation and training are affordable.	
ID	CRITERIA	TARGET THRESHOLDS
5.1.1	Cost of initial take-up and continued use is reasonable and feasible.	— Cost of training and implementation is deemed reasonable by a panel of program and systems leaders based on cost study, relative to the typical program budget.
5.1.2	Observation tool does not require expensive equipment purchases.	— Platform for training, data entry, and reporting leverages commonly available technology (e.g., standard laptops, cell phones, and standard video and digital audio files).
SUBGOAL 5.2	Observation tool is convenient to use regularly in publicly funded pre-K systems and centers.	
ID	CRITERIA	TARGET THRESHOLDS
5.2.1	Observation tool requires a low to moderate amount of training for reliable use by persons with at least 2 years of experience in early childhood education classrooms.	— Both initial and refresher trainings include a mix of synchronous and asynchronous training sessions, with an initial training of as low as 15 minutes for more straightforward assessment modules and up to 1 hour for more complex modules. For example, observation tools that focus on indicators of practices that are easy to determine will enable faster training and sufficient reliability. Special modules that go deeper into specific delivery of practices may require more training time. — Training for program administrators on how to do instructional coaching observations should be available through a variety of outputs (online, in person, hybrid) and through a train-the-trainer option (e.g., teacher-to-teacher, expert-teacher). — Observation tool documentation and training materials are provided in multiple formats and are accessible to persons who do not have technical expertise, including video examples to demonstrate procedures. — Training and administration can be modularized by individual subdomains and therefore be shorter (e.g., less than an hour). — Documentation and data collection forms are available in languages of instruction and written at or below eighth-grade reading proficiency, with hyperlinks to explain technical terms.

Goal 5: Scalability

ID	CRITERIA	TARGET THRESHOLDS
5.2.2	Observation tool or subscales of observation tool can be completed within a short period (less than an hour).	— Data entry and reports support the submission and use of data for only one or some subscales and domains. — Technology-enabled options, if used, reduce the time required to record and review observations with peer teachers, coaches, or program directors.
5.2.3	Analyzing and summarizing observed results is easy and can be done in a reasonable amount of time.	— Preparing a summary of observed results requires, on average, less than 30 minutes for each observation or observation session. — 80% of observers report that guidelines and interface for data entry are easily implementable.
5.2.4	Data entry is easy and can be done in a reasonable amount of time using typical IT infrastructures of pre-K programs.	— Data entry requires, on average, less than 30 minutes for each observation session. — Data entry requires no more than a laptop/computer and/or cell phone. — Data entry can be completed when Wi-Fi is available and can also be conducted asynchronously in offline mode, as needed, with data capture when Wi-fi is available. 80% of observers report that guidelines and interface for data entry are easily implementable. — Observer instructions for setup are in English but are also available in Spanish for data entry when an observation session is completed in Spanish. Instructions in other languages of instruction are encouraged when possible.
SUBGOAL 5.3	Platforms for training, data storage, and reporting ensure data security and support data interoperability.	
ID	CRITERIA	TARGET THRESHOLDS
5.3.1	Data (e.g., ratings, notes, videos) are securely uploaded and stored.	— Technical parameters and processes for data systems and data management infrastructures ensure the privacy and security of information collected and stored about children and teachers. — Platform maintains and stores data about children in compliance with requirements of the Family Educational Rights and Privacy Act (FERPA), Children's Online Privacy Protection Act (COPPA), Individuals with Disabilities Education Act (IDEA), and Children's Internet Protection Act (CIPA).

Goal 5: Scalability

ID	CRITERIA	TARGET THRESHOLDS
5.3.2	Observation tool platform is interoperable with other commonly used classroom data platforms.	— Observation tool system has easy data export functions. — Data from the observation tool can be integrated with data from other classroom-level observation tools, child assessment data, and curricula.
5.3.3	Updates to data structures in existing observation tool platform can be made without loss of collected information about prior observations.	— Updates to data dashboards with prior data and reports maintain past data and legacy reports. — Prior data are automatically uploaded into new reports to support longitudinal tracking. — When changes to how data are collected or displayed are made, users receive guidance on how to interpret findings.

Acknowledgments

Thank you to the families, educators, coaches, administrators, and early childhood stakeholders from over 20 states and localities who advised this work and took the time to openly share their hopes and dreams for young children, their experiences and perspectives with assessments, and their lived expertise more broadly. The importance of continuing to integrate feedback and engage in co-design activities with these priority stakeholders, whose voices are often not heard, is a central theme throughout this document.

Additionally, an Advisory Group was convened by the SRI Education and Substantial teams at multiple points during our writing process (see names at right). We appreciate their time and willingness to provide their insights about how observations can be designed and utilized to support the needs of Pre-K educators, families, and young children from diverse, multilingual, and multi-racial backgrounds. Blending policy and research considerations can be challenging, but their openness to explore both the complexities of this space, as well as voice potential solutions, greatly informed the writing of this document. It should be noted, however, that this document does not necessarily reflect their practices or points of view. We also appreciate the reviews and comments provided by staff and consultants of Teachstone, University of North Carolina Frank Porter Graham Child Development Institute, and the Branagh Group based on their experience developing and implementing tools that assess teaching and classroom quality. Their reviews helped us to clarify the descriptions of the criteria and thresholds and focus on factors within the control of coaching tool developers.

This resource was co-developed by Krystal Thomas, Todd Grindal, Daisy Rutstein, Gullnar Syed, Sarah Gerard, Shari Golan (SRI), and Sheryl Cababa, Amanda Di Dio, Behnosh Najafi, and Kat Ward (Substantial).

Photos by Evan McGlinn, Mark Makela, and Allison Shelley/The Verbatim Agency for EDUimages.

Icons created by DinsosoftLab, Adrien Coquet, and Atif Arshad from the Noun Project.

Advisory Group Members:

Dr. Jennifer Keys Adair

University of Texas at Austin

Dr. Courtney Bell

University of Wisconsin–Madison

Dr. Lillian Duran

University of Oregon

Dr. Tonia Durden

Georgia State University

Dr. Caroline Ebanks

U.S. Department of Education
Institute of Education Sciences
National Center for Education Research

Dr. Bryant Jensen

Brigham Young University

Victoria Jones

National Head Start Association

Sara Mead

DC Office of State Superintendent
of Education, Early Learning

Endnotes

- ¹ Yoshikawa, H., Weiland, C., & Brooks-Gunn, J. (2016). When does preschool matter? *The Future of Children*, 26(2), 21—35. <https://eric.ed.gov/?id=EJ1118535>
- ² Our definition of “high-quality” does not include the commonly used term “developmentally appropriate.” We refrain from using this term as it does not pay sufficient attention to the culturally and linguistically responsive needs and practices of children and their unique and distinct skills. Instead, we encourage engagement in a variety of activities to support children’s skills and growth while elevating home language and culture.
- ³ Walters, C. R. (2015). Inputs in the production of early childhood human capital: Evidence from Head Start. *American Economic Journal: Applied Economics*, 7(4), 76—02. <https://doi.org/10.1257/app.20140184>
- ⁴ Early, D. M., Maxwell, K. L., Ponder, B. D., & Pan, Y. (2017). Improving teacher–child interactions: A randomized controlled trial of Making the Most of Classroom Interactions and My Teaching Partner professional development models. *Early Childhood Research Quarterly*, 38, 57—70.; Isner, T., Tout, K., Zaslow, M., Soli, M., Quinn, K., Rothenberg, L., & Burkhauser, M. (2011). Coaching in early care and education programs and quality rating and improvement systems (QRIS): Identifying promising features. *Child Trends*.
- ⁵ Here, we use the phrase instructional coach, but coach can be many different roles, including a dedicated teacher-coach, a program administrator, a peer teacher, or other individual supporting teachers in their classroom practices.
- ⁶ e.g., Curenton, S. M., Iruka, I. U., Humphries, M., Jensen, B., Durden, T., Rochester, S. E., Sims, J., Whittaker, J. V., & Kinzie, M. B. (2020). Validity for the Assessing Classroom Sociocultural Equity Scale (ACES) in early childhood classrooms. *Early Education and Development*, 31(2), 284—303.; Gordon, R. A., Fujimoto, K., Kaestner, R., Korenman, S., & Abner, K. (2013). An assessment of the validity of the ECERS–R with implications for measures of child care quality and relations to child development. *Developmental Psychology*, 49(1), 146—160. <https://doi.org/10.1037/a0027899>; López, F. (2011). The nongeneralizability of classroom dynamics as predictors of achievement for Hispanic students in upper elementary grades. *Hispanic Journal of Behavioral Sciences*, 33(3), 350—376. <https://doi.org/10.1177/0739986311415222>
- ⁷ Consistent with experts in the field, this document uses the gender-neutral term “Latine” to refer to individuals whose cultural background originated in Latin America. In U.S. academic circles, “Latinx” is often used as a gender-inclusive term to refer to people from Latin American backgrounds, but Spanish-speakers find that Latinx is unpronounceable in Spanish. Therefore, we have opted to use the term Latine, which is commonly used throughout Spanish-speaking Latin American countries (see Melzi et al., 2022).
- ⁸ The term “multilingual children” refers to youth, beginning at birth and continuing through college and into the workforce, who learn English while speaking one or more other languages.

Endnotes (continued)

- ⁹ We use “teachers” as shorthand for all classroom staff, which also include teacher assistants, aides, and paraprofessionals, who may be part of the observation process.
- ¹⁰ MDRC and Substantial, *User-Informed Principles: Developing Assessments for All Early Learners* (New York: MDRC, 2022).
- ¹¹ While prekindergarten-age children receive care in a variety of settings, such as family and home-based care, this TPP is focused on publicly funded center-based settings because of the current federal investments in pre-K expansion and improvement. Further, to provide sufficiently detailed guidance about the content and uses of the observation tool, the TPP focuses on pre-K rather than trying to cover the entire early learning age range (birth through kindergarten).
- ¹² These prioritized populations are not a monolith, and we expect that considerable heterogeneity will exist even within these groups, resulting from the many ways in which these categories might intersect (e.g., across race and income, or those who are monolingual speakers vs. those who are multilingual speakers). These intersections will have important implications for how assessment items and solutions are constructed. In addition, poverty should be addressed as a structural barrier that can also manifest differently across ecologies (e.g., immigration status, intergenerational poverty, rural and urban poverty). Future assessments will need to account for both variation and commonalities in how learning and developmental constructs appear across children’s varied social and cultural contexts.
- ¹³ Rimm-Kaufman, S.E., & Thomas, K. (2021). How White, middle class teachers can apply psychology to teach students who are different from them (Practice Brief Vol. 2, No. 2). American Psychological Association Division 15 Educational Psychology. <https://apadiv15.org/wp-content/uploads/2022/08/Practice-Brief-Rimm-Kaufman-Thomas.pdf>
- ¹⁴ Brooks, J. L., Gayl, C. L., & Wernstedt-Lynch, C. (2022.) *Measuring the quality of early learning environments: A guide to evaluating ideal learning environments for young children*. Trust for Learning.
- ¹⁵ Iruka, I. U., Curenton, S. M., & Eke, W. A. I. (2014). *The CRAF-E4 family engagement model: Building practitioner's competence to work with diverse families*. Elsevier Academic Press.; National Association for the Education of Young Children. (2019). *Advancing equity in early childhood education*. <https://www.naeyc.org/sites/default/files/globally-shared/downloads/PDFs/resources/position-statements/advancingequitypositionstatement.pdf>
- ¹⁶ National Academies of Sciences, Engineering, and Medicine. (2017). *Promoting the educational success of children and youth learning English: Promising futures*. The National Academies Press. <https://doi.org/10.17226/24677>
- ¹⁷ Hehir, T., & Katzman, L. I. (2012). *Effective inclusive schools: Designing successful schoolwide programs*. John Wiley & Sons.
- ¹⁸ American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. AERA.
- ¹⁹ Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33(1), 159—174.



Thank you

BILL & MELINDA
GATES foundation

SRI Education™
A DIVISION OF SRI INTERNATIONAL

substantial

Thomas, K., Grindal, T., Rutstein, D., Syed, G., Gerard, S.N., Golan, S. (2023). *Reimagining Instructional Coaching: Developing Observation Tools to Support Instructional Coaching in Pre-K Classrooms*. SRI International.